

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2001, 16:20:13 ; Search time 7577.75 Seconds  
(without alignments)  
11971.674 Million cell updates/sec

Title: US-09-698-903B-1  
Perfect score: 5865  
Sequence: 1 aattacaacggtatatatcc.....caattgaatatatctgcgcg 5865

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues  
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_ba3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_om:\*  
8: gb\_ov:\*  
9: gb\_pat1:\*  
10: gb\_pat2:\*  
11: gb\_ph:\*  
12: gb\_pl1:\*  
13: gb\_pl2:\*  
14: gb\_pl3:\*  
15: gb\_pl4:\*  
16: em\_ba1:\*  
17: em\_ba2:\*  
18: em\_fun:\*  
19: em\_htgo\_hum:\*  
20: em\_htgo\_inv:\*  
21: em\_htgo\_rod:\*  
22: em\_htg\_hum1:\*  
23: em\_htg\_hum2:\*  
24: em\_htg\_hum3:\*  
25: em\_htg\_hum4:\*  
26: em\_htg\_hum5:\*  
27: em\_htg\_hum6:\*  
28: em\_htg\_hum7:\*  
29: em\_htg\_hum8:\*  
30: em\_htg\_inv1:\*  
31: em\_htg\_inv2:\*  
32: em\_htg\_other:\*  
33: em\_htg\_rod:\*  
34: em\_hum1:\*  
35: em\_hum2:\*  
36: em\_hum3:\*  
37: em\_hum4:\*  
38: em\_hum5:\*  
39: em\_hum6:\*  
40: em\_hum7:\*  
41: em\_in:\*  
42: em\_om:\*  
43: em\_or:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_ph:\*  
47: em\_pl:\*  
48: em\_ro:\*  
49: em\_sts:\*  
50: em\_sy:\*  
51: em\_un:\*  
52: em\_vi:\*  
53: gb\_sts1:\*  
54: gb\_sts2:\*  
55: gb\_sts3:\*  
56: gb\_sy:\*  
57: gb\_un:\*  
58: gb\_vil:\*  
59: gb\_vil2:\*  
60: gb\_vil3:\*  
61: gb\_vil4:\*  
62: gb\_vil5:\*  
63: gb\_vil6:\*  
64: gb\_vil7:\*  
65: gb\_vil8:\*  
66: gb\_vil9:\*  
67: gb\_vil10:\*  
68: gb\_vil11:\*  
69: gb\_vil12:\*  
70: gb\_vil13:\*  
71: gb\_vil14:\*  
72: gb\_vil15:\*  
73: gb\_vil16:\*  
74: gb\_vil17:\*  
75: gb\_vil18:\*  
76: gb\_vil19:\*  
77: gb\_vil20:\*  
78: gb\_vil21:\*  
79: gb\_vil22:\*  
80: gb\_vil23:\*  
81: gb\_vil24:\*  
82: gb\_vil25:\*  
83: gb\_vil26:\*  
84: gb\_vil27:\*  
85: gb\_vil28:\*  
86: gb\_vil29:\*  
87: gb\_vil30:\*  
88: gb\_vil31:\*  
89: gb\_vil32:\*  
90: gb\_vil33:\*  
91: gb\_vil34:\*  
92: gb\_vil35:\*  
93: gb\_vil36:\*  
94: gb\_vil37:\*  
95: gb\_vil38:\*  
96: gb\_vil39:\*  
97: gb\_vil40:\*  
98: em\_ba3:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match %	Length	DB ID	Description
1	4903.2	83.6	4946	9	A60108	A60108 Sequence 1
2	4903.2	83.6	4946	9	A76915	A76915 Sequence 1
3	4903.2	83.6	4946	9	AR098307	AR098307 Sequence 1
4	2386.8	40.7	5560	9	A60112	A60112 Sequence 5
5	2386.8	40.7	5560	9	AR098311	AR098311 Sequence
6	1736.6	29.6	90387	12	AC007152	AC007152 Arabidops
7	1736.6	29.6	101933	12	AC004146	AC004146 Arabidops
8	1732.4	29.5	2821	13	ATRCBSA	X13611 Arabidops1s





Qy	3361	ggtaaccatgtagtaactaattcttcttaagtaaaaaacitltagtttgagtgatgtgtgtact	3420
Db	3360	GGTACCATTGTTAGCTAAATTCCTTTAAGTAAAAACCTTTGATTTTGAGTGATGATGTTGTACT	3419
Qy	3421	gttaaccttgcacacacaaaggccataatagagcacaagaacacatacacacaacacttgcacaaa	3480
Db	3420	GTTACACTTGCACCAAGGGCATAATAGAGCACAAGACATACACAACAACCTTGCAAAA	3479
Qy	3481	ctaaactttgttgaggacatttcgaggaaaaatggggagtagcagcgttaactcagagggtaac	3540
Db	3480	CTAACTTTTGTGGAGCATTTGAGGAAATGGGAGTAGCAGGCTAATCTGAGGGTAAC	3539
Qy	3541	attaagggttcacatgtattaatttggcacaacatggacctagtgtaggagaaaaaagtagacca	3600
Db	3540	ATTAAAGGTTTCATGTATTAATTTGTTGCAACATGGACTTAGCTGAGGAAAAAGTAGCCA	3599
Qy	3601	aaatttggctacacctgatttcagttatcggaataacatatatgaagctgtcgtagagaa	3660
Db	3600	AAATTTTGTCTCACCCTGATTTTCAGTTATGGAATTCAGATTTATGAAGCTGTGCTAGAGAA	3659
Qy	3661	gaagtttactctagtcagccacccaccttatgcgaagtcctgcttttagcttgatgcacaaa	3720
Db	3660	GATGTTTATTCAGTCCAGCCACCCTTATGCAAGTCTGCTTTTACGTTGATTCACAAA	3719
Qy	3721	actgatttaattcacatttgctaataatgtgcatacttcggacctatgtcgcctttaactcogag	3780
Db	3720	ACTGATTTAATTTACATTTGCTAAATGTGCATTACTTTCGAGCCTATGTGCTTTAATTTCGAG	3779
Qy	3781	tagagtcataatatagtagcacataaaaaatcagtttgatgaatcatcttcaataaagtgacaa	3840
Db	3780	TAGATGCTATATATTAGTACATAAAAAATCATGTTTGAATCATCTTTCATAAAGTGACAA	3839
Qy	3841	gtcaattgctccctctgttttggcactatattccaactctgtttaatgcgaataattcacagtta	3900
Db	3840	GTCAATTGTCCCTTCTGTTTGGCACTATATTCAAATCTGTTAATGCAAAATTATCCAGTTA	3899
Qy	3901	tacttagctagatatccaattttgataaaaaatagctcttgattagtaaacccggatagtg	3960
Db	3900	TACTTAGCTAGATATCCAAATTTTGAATAAAAATAGCTCTTGATTAGTAAACCCGGATAGTG	3959
Qy	3961	acaaagtccacataccatacceaactctggctgctggctgaagttcttgatcgacatgggg	4020
Db	3960	ACAAAGTCACATATCCATCAAACTTCTGGTGCTGCTGCTTAAGTTCGTGACATCGGG	4019
Qy	4021	ttaaaatttaaatgggacacataaatagcctatttggcaaaatctcccacatcgaaaaatg	4080
Db	4020	TTAAATTTAAATGGGACACATAAATAGCTATTGTCGCAAAATCTCCCCATCGAAAATG	4079
Qy	4081	acagatgtgtacatggaacacaaaagtcctctgatagaagtcgcaagatcacacaattt	4140
Db	4080	ACAGATTTGTATCGAAAACAAAAAGTCTCTGTATCAAGTCGCAAGTATCACAATTT	4139
Qy	4141	tcctatcgagagatagattgaagaagtcaggggagaagcgttaactggaacataacacaat	4200
Db	4140	TCATCGAGAGATAGAAATTTATGCTCCAAAATAAGTGGAGCGGTTAACTGGAACATAACACAA	4199
Qy	4201	gtctaaattcaattgcattcgtcaacacaaaagtgtattcaactctccgggtccacaataag	4260
Db	4200	GTCTAAATTAATTGCAATTCGCTTAACCAAAAAGTGATATTACTTCTCTCCGGTCCACAATTA	4259
Qy	4261	ttattttttggocctttttttatggtccaaaaataagtgagtttttttagatttcaaaaaatg	4320
Db	4260	TTATTTTTTTGGCCCTTTTTTTATGCTCCAAAATAAGTGAGTGTGAGTTTTCAGATTTCAAAA	4319
Qy	4321	attaactatttttttactcacagtcacctggagctaaatgctgttgagtatgtattgtaga	4380
Db	4320	ATTATATATTTTTTTTACATACAGTGCCTTGAGGTAAATGGTGTGGAGTATGTGTGTAGA	4379
Qy	4381	aatgtttatgtgaagaaatagtaaaaggttaatatgatccaatttcaattgcatttaattgttt	4440
Db	4380	AAATGTTTATGTGAAGAAATAGTAAAGGTTTAATATGATCAATTTTCATTTGCTATTAATGTTT	4439

QY	4441	aaaaatgtgaattcttaatctgtgtgaataaacacacaaataacattatgttgacacgga	4500
Db	4440	AAAAATGTGAATTTCTTAATCTGTGTGAAAAACACCAAAAAATCACTTATTGTGACCGGA	4499
QY	4501	gaaagtacataaaatataatttggaagcgactaaaaataaaacttttccatattatcacga	4560
Db	4500	GAAAGTATATAAATATATATTTGGAGCGACTAAAAATAAACTTTTCATATTATTACGA	4559
QY	4561	acctaaaaacagcatatgtgagttctctagggaacttaaatcactaaaaataataataatttaac	4620
Db	4560	ACCTTAAAAACAGCATATGGTGTAGTTCTTAGGGAATCTAAATCACATAAAATTAATAAAGAA	4619
QY	4621	gcacacagatcaatacatatgatttacacgctcaaacacgaaattcgttaaatatttaac	4680
Db	4620	GCACACAGTATCAATACATATGATTACACCGCTCAACACGAAATTCGTAAATATTTAAT	4679
QY	4681	ataataagaataatccaaatagcctcccaccctataaacttaaaactaaaaataaacacgc	4740
Db	4680	ATAATAAAGAATTAATCCAAATAGCTCCACCCCTATAAATTAACATAAAAAATAACACG	4739
QY	4741	gaatgtatatatgcataatttatatatataattaaatgtgtataataatgcataataatgc	4800
Db	4740	GAATGTATATTATATGCAATAATTATATTTAAATGTGTATTAATCATGTATAATCAATGT	4799
QY	4801	ataactctgtatctggttagaataaagtaaacataataatagccgctatttgtctaa	4860
Db	4800	ATAATCTATGTATGTTAGAAAAAGTAACAATTAATATAGCCGGCTATTGTGTGTA	4859
QY	4861	aatccctaataatcgcgcgagatccccgggaattccggggaagcttagatccatgcag	4920
Db	4860	AATCCCTAATAATATCGGACGGATCCCGGGAAATCCGGGGAAGCTTAGATCCATGGAG	4919
RESULT	2		
A76915			
LOCUS	A76915	4946 bp	DNA
DEFINITION	Sequence 1 from Patent EP0757102.		PAT
ACCESSION	A76915		
VERSION	A76915.1	GI:6088712	
KEYWORDS			
SOURCE	Transformation vector pTHW107.		
ORGANISM	Transformation vector pTHW107		
REFERENCE	artificial sequence; vectors.		
AUTHORS	De,B.N.		
TITLE	GENETIC TRANSFORMATION USING A PARP INHIBITOR		
JOURNAL	Patent: EP 0757102-A 1 05-FEB-1997;		
FEATURES	PLANT GENETIC SYSTEMS NV (BE)		
	Location/Qualifiers		
	1..4946		
	/organism="Transformation vector pTHW107"		
	/db_xref="taxon:126810"		
BASE COUNT	1569 a	891 c	963 g
ORIGIN	1523 t		
Query Match	83.6%	Score 4903.2;	DB 9; Length 4946;
Best Local Similarity	99.9%	Pred. No. 0;	
Matches 4916;	Conservative	0; Mismatches	3; Indels 1; Gaps 1;
QY	1	aattacaacggtatatatcctgccagctactcgcgcgtcgaactcgcgcgtcgagacatg	60
Db	1	AAATACACGGTATATATCTGCCAGTACTCGGCCGTCGAACTCGGCCGTCGAGTACATG	60
QY	61	gtcgataagaagaaggcaatttggtagtggtaattcccatctctgaaagaataatagtttaa	120
Db	61	GTGCATAAGAAAAGGCAATTTGTAGATTAAATCCCATCTTGAAGCAAAATATAGTTTAA	120
QY	121	atatttattgataaaaaacaagccaagttattatagtcgaagcaaaaaataatttatt	180
Db	121	ATATTTATGTATAAAATAACAAAGTCAGGTATATAGTCCAAAGCAAAAACATAAATTTATT	180
QY	181	gatgcaggtttaaactcagaataatttcaataactgatttatcatcagctggtacattgcgc	240

Db	1260	TTTACTTACTATAGAGCTTTCATACCTTTTTTTTTTTTACCTTGGATTAGTTTAATATATAATG	13131
Qy	1321	gtatgattcatgaataaaaaatggaaaattttttgaattttgtactgctaaatgcataaagatt	1380
Db	1320	GTATGATTCTAAGAATAAAAAATGGAAATTTTTTGAATTTGTACTGCTAAATGCATAAGATT	1379
Qy	1381	aggtgaaaactgtgggaatatatatatttttttttcattttaaagcaaaaatttgccttttactaga	1440
Db	1380	AGGTGAACCTGTGGGAATATATATTTTTTTTCATTTAAAAAGCAAAATTTGCCCTTTTACTAGA	1439
Qy	1441	attataaaatagaaaaatatataacattccaataaaaaatgaaaatagaagactttcaaaa	1500
Db	1440	ATTATAAATAATAGAAAAAATATATAACATTCAAAATAAAAATGAAAAATAGAAACTTTCAAAA	1499
Qy	1501	aacagaactatgtttaatgttaaaagttagtcgcacatcaaagtcatctctgtttacaatatg	1560
Db	1500	AACAGAACTATGTTTAAATGTGTAAGATTTAGTCGCACATCAAGTCTATCTGTTACAATATG	1559
Qy	1561	ttacaacaagtcatagccccaacaaagtttagcacgtctaaataaaactaaagaagtcacaga	1620
Db	1560	TTACAACAAGTCTAAGCCCAACAAAGTTTAGCACGTCTAAATAAACTAAAGAGTCCACGA	1619
Qy	1621	aaattaccaaaataaagcccaacaaaagttattgaataaaaaaaaacacgcaccaaca	1680
Db	1620	AAATATTACAATATCATAGCCCAACAAAGTTATTGATCAAAAAAAAAGGCCACACA	1679
Qy	1681	agctaaacaaagtcctcaaaaaaactctccaagtcctcatcttcctttatgaacattgaaa	1740
Db	1680	AGCTAAACAAGTCCAATAAAAAAACTTCTCAAGTCTCCATCTTCCTTTATGAACATTGAAA	1739
Qy	1741	actatacacaacaaagtcagataaaatctctcttctgggcgtgtcttcccaacctcttaca	1800
Db	1740	ACTATACAAAAACAAGTCAGATAAAATCTCTTCTTGGGCGCTGTCCTCGCAACCTCCTACA	1799
Qy	1801	tcactctcctatcgagttgaatgttttactctataccttttccgtttgcgaatgaattgata	1860
Db	1800	TCACTTCCCCTATCGGATTTGAATGTTTTTACTTGTACTCTTTCGGTTTGCATGATATTGATA	1859
Qy	1861	gtatgtttgtgaaaactaatlagggttcaacaatcgaaagtcattggaatatggatttgggtcca	1920
Db	1860	GTATGTTTGTGAANAATAATAGGGTTTAACAAATCGAAGTCATGGGAATATGGATTTGGTCCA	1919
Qy	1921	agattttccgagagcttctctagtagaaaagcccatcaccagaaaatttactagtaaaataaaa	1980
Db	1920	AGATTTTCCGAGAGCTTCTCTAGTAGAAAAGCCCATCACCAGAAATTTTACTAGTAAAAATAAA	1979
Qy	1981	tcaccaatagggtttcttattatgtgcgaaaattccaataataattatagaggatatttcaaaa	2040
Db	1980	TCACCAATTAGGTTCTTATTATGTGCCAAATTCAAATAATTAATATAGAGGATATTTCAAA	2039
Qy	2041	tgaaacgctatgaatgttattagtaaaatgggtcaggttaagacattaaaaaaatccttcagtc	2100
Db	2040	TGAACAACGTATGAATGTTATTAGTAATAATGGTCAGGTAGACATTTAAAAAATCCTACGTC	2099
Qy	2101	agatttcaactttaaaaattcgatcagttggaattgtcaaaaaatttgggatcttacta	2160
Db	2100	AGATATTCAACTTTAAAAATTCGATCAGTGTGGAATTTGTACAAAAATTTTGGGATCTACTA	2159
Qy	2161	tatatataaattgctttacaacaacttggatttttttttggaggctgggaatttttaactta	2220
Db	2160	TATATATATAATGCTTTTACAACTTTGGAATTTTTTTTTTGGAGGGCTGGAAATTTTAAATCTA	2219
Qy	2221	cataattgttttggccatgcaccaactcatgttttagtgttaactcttggatttggttgcga	2280
Db	2220	CATATTGTTTTTGGCCATGCACCACCACTCAATGTTTAGTGTGAATACTTTGATTTTGTCCAA	2279
Qy	2281	tatatgttctgtgtatatttgtataagaatttcttttgaccatacacacacacatatata	2340
Db	2280	TATATGTGTCGTGTATATTGTATTAAGAAATTTCTTTGACCATATACACACACATATA	2339
Qy	2341	tatatatatatatatatatatcatgcaccttcttaattgaaaaaataatatatat	2400
Db	2340	TATATATATATATATATATATCATGCACCTTTTAAATTTGAAAAAATAATATATATATAT	2399

QY 2401 atagtgcaatttttttcaacaaccatataatgttgcgattgactgcacaaataactgctaga 2460  
Db ATAGTGCAATTTTCTACAAACCATATATGTTGGATTGATCTCAAAAATACATGCTAGA 2459  
QY 2461 gtaatgaaaaataatactatgtgtgaataattatctcagatgtttaagattttcttaagta 2520  
Db GTAAATGAAAAATAAATCTATTGCTGAAATATATCTCAGATGTTAAGATTTCTTTAAAGTA 2519  
QY 2521 aattctttcaaaatttttagctcaaaagctctgtataaactaaagaataatacacaaatctcga 2580  
Db AATTTCTTTCAAAATTTAGCTAAAGAGTCTTGTAATAACTAAGAAATAATACACAATCTCGA 2579  
QY 2581 ccacggaaaaaaacacataataaatttgaaatttcgaccgctgaccccggaattcgaact 2640  
Db CCACGGAAAAAAACACATAATAATTTGAATTTGACCGCGGTACCCGGAATTCGAGCT 2639  
QY 2641 cggtaaccggggactctcccgatcagatagaaacatagatgacacccgcgcgataattatc 2700  
Db CGGTACCCGGGGATCTCCCGATCTTAGTAACATAGATGACACCGCGCGGATAAATTTATC 2699  
QY 2701 ctgatttcgcgcgtatattttgtttctatcgcgatttaaatgtataaattcgaggactct 2760  
Db CTAGTTTGCgcgtATATTTGTTTCTATCgcgtATTAATGTATTAATTTGCGGGACTCT 2759  
QY 2761 aatcataaaaaaccatctcotaataacgtcatcgtcattacatgttaattattacatgctt 2820  
Db AATCATAAAAAACCATCTCATATAATACGTCATGCATTTACATGTTAATTTACATGCTT 2819  
QY 2821 aacgttaattcaacgaataattatgataatcatcgcgaagaccgcgcgaaggattcgaatct 2880  
Db AACGTAATTCACAGAAATTTATGATAATCATCGGAAGACCGGCAACAGGATTCATCT 2879  
QY 2881 taagaaactttatgccaatgtttgaacgatctgtctcgatctctctagacgcggaag 2940  
Db TAAGAACTTTATTGCAAAATGTTTGAACGATCTGCTCGGATCCCTAGAGCCGGAAG 2939  
QY 2941 tgaattgcgcgacagagtttgaaagaaaaatttatcacacatttatgaaagctgaaa 3000  
Db TGAATTGACCGATCAGAGTTTGAAGAAAAATTTATTACACACTTTATGTAAGCTGAAA 2999  
QY 3001 aaacggcctccgcaggaagcgcttttttctgttatctgtatttttgaagctctgataa 3060  
Db AAAACGGCTCCGCAGGAAGCGCTTTTTCGTTATCTGATTTTGTAAAGCTCTGATAA 3059  
QY 3061 tggctcgtgttttgaatacgcagcagctgcgtgtgagtaagaatccggtctgaaattctg 3120  
Db TGGTCCGTTGTTTGTAAATCAGCCAGTCGCTTGAGTAAGAATCCGGTCTGAATTTCTG 3119  
QY 3121 aegcctgatgtatagtttaataatccgcttcacgcgatgttcctcgcttttgcccgaggat 3180  
Db AAGCCTGATGATAGTTAATATCCGCTTCACGCCATGTTCTGCTCCGCTTTTGCCCGGAGT 3179  
QY 3181 tbgcttccctgtttgagaagatgtctccgcgatgcttttcccggaagcgcgctctgca 3240  
Db TTGCTCTCCCTGTTTGAAGAATGTTCTCCGCGATGCTTTTCCCGGAGCGCTGCA 3239  
QY 3241 aggttcccttttgatgcacccagcagggctgtgtctctgattttgtgaatgtaata 3300  
Db AGGTTCCCTTTTGATGCCACCCAGCGAGGGCTTGTCTCTGATTTTGTAAATGTAATA 3299  
QY 3301 tcaggtagcttatgatgtctgaagataataccgcgaaccccgatcaaacggtgtgataacc 3360  
Db TCAGGTAGCTTATGATATGCTCTGAAGATAATCCGCAACCCCGTCAACCGTGTGATAACC 3359  
QY 3361 ggtacatggttagctcaatttcttaagttaaaactttgatttgatgtatgttgtact 3420  
Db GTACCATGGTAGCTAAATTTCTTTAAGTAAAAACTTTTGAATTTGATGTGATGTTGTACT 3419  
QY 3421 gttacacttgcaccacaagggcataatagagcaacaacatacaacaacacttgcgaaaa 3480  
Db GTTACACTTGCCACCACAGGGGCATATATAGACCAACAGACATACACAACACTTGCACAAA 3479

QY 3481 ctaacttttctgagcatttcgaggaataatggggagtagcagcgtacatactctgagggtaac 3540  
Db CTAACTTTTCTGGAGCATTTTCGAGAAAAATGGGAGTAGCAGGCTAATCTGAGGGTAAC 3539  
QY 3541 attaaagtttcacatgtatttaatttggcaaacatggaacttagtgtgaggaataaagtacca 3600  
Db ATTAAGGTTTCATGTATTAATTTGTTGCAACATGGACCTAGTGTGAGAAAAAAGTACCA 3599  
QY 3601 aaatttgcctccacctgatttcagttatggaataacacattatgaagctgtgctagagaa 3660  
Db AAATTTTGTCTCACCCTGATTTTCAGTTATGGAATTTACATTTATGAAGCTGTCTAGAGAA 3659  
QY 3661 gatgtttattctagtcagcccccacccttatgcaagtcgtgttttagcttatttcacaaa 3720  
Db GATGTTTATTTCTAGTCCAGCCACCCACCTTATGCAAGTCTGCTTTTGTAGCTTATTTCA 3719  
QY 3721 actgatttaatttacattgtctaaatgctacattctcagccctatgctgcttatttcagag 3780  
Db ACTGATTTAATTTACATTTGCTTAAATGTCATCTCTGAGCCTATGTCGCTTTAATTCGAG 3779  
QY 3781 taggatgtatatattagtagacataaaaaatcatgttttgaatcatctttcataaagtacaa 3840  
Db TAGGATGTATATATTAGTACATAAAAAATCATGTTGAATCATCTTTTCATAAAGTGACAA 3839  
QY 3841 gtcaattgtccctctctgtttgcaactatattcaatctgttaatgcacaaattatccagtta 3900  
Db GTCAATTGTCCTCTGTTTGGCACTATTTCATCTGTTAATGCAATTTATTCAGATTA 3899  
QY 3901 tacttagctagatatccaaattttgaaataaaaatagctctctgtattagtaaacccggatagt 3960  
Db TACTTAGCTAGATATCCAAATTTTGAATAAAAAATAGCTCTTGATTAGTAACCGGATAGT 3959  
QY 3961 acaaaagtcacatcatccatacaaaactctggctcgtcgtgaagtctgatgcacatgggg 4020  
Db ACAAGTCACATATCCATCAAACTTCTGGTCTGCTGCTAGTCTGATGACATGGG 4019  
QY 4021 ttaaaattttaaattgggacacataaaatagcctatttgcacaaatctcccatcgaataatg 4080  
Db TTTAAATTTAAATTTGGGACACATAAATAGCCTATTGTTGCAAAATCTCCCCATCGAAAATG 4079  
QY 4081 acagattgttacatgggaacacaaaaagtcctctgtatagaagtcgcaaatgatacaattt 4140  
Db ACAGATTGTTATCGAAGAAAAACAAAAGTCTCTGTATAGAAGTCGCAAAAGTATCACAATTT 4139  
QY 4141 tctatcagagatagatgtgaagaagtcagggagcgttaactggaacacatacaacat 4200  
Db TCTATCGAGAGATAGATTGAAGAAGTGCAGGAAGCGGTTAACTGGGAACATAACACAAAT 4199  
QY 4201 gtcataaataattgcatcgttaaccacaaaaagtgattactctctccggtccacaaataag 4260  
Db GTCTAAATTAATTTGATTCGCTAACCAAAAAGTGTATTACTCTCTCCGGTCCACAATAAG 4259  
QY 4261 ttatttttggcccttttttatttggtccaaaataagtgagtttttttagatttccaaaatg 4320  
Db TATTTTGTGGCCCTTTTATGCTCCAAAATAAGTGAGTGTGTTTATGATTTCAAAAATG 4319  
QY 4321 atttaattattttttactacagtcgcttgagtaaatggtgtgagtatgtgttaga 4380  
Db ATTTAATTAATTTTACTACAGTCCCTTGGAGTAAATGCTGTTGGAGTATGTGTAGA 4379  
QY 4381 aatgtttatggaagaaatagtaaggttaataatgataatctcatctgctatttattt 4440  
Db AATGTTTATGTGAAGAAATAGTAAGGTTAATAATGATCAATTTTCATTTGCTATTTAATGTT 4439  
QY 4441 aaaatgtgaatttcttaactctgtgaaacacacaaaaatacacttatgttgaccgga 4500  
Db AAAATGTGAATTTCTTAATCTGTGAAAAACACCAAAAATCACTTATTGTGACCCGGA 4499  
QY 4501 gaaagtataataatatabatttggaaagcgaactaaaaataaacttttctcatattacga 4560  
Db GAAAGTATATAAATATATATTGGAAGCGACTAAAAATAAACTTTTCTCATATTATACGA 4559  
QY 4561 acctaaaaacagcatatggtagtttcttagggaatctaaactacataaataaagaa 4620









Qy	4741	gaatgtatattatcatgataaattatattatataattataataatgtgtataataatcatgataataatcattgt	4800
Db	4740	GAATGTATATTATATATGCATAAATTTATATATATTAATATGTTATTAATCATGTATATAATCAATGT	4799
Qy	4801	ataatctatgtatgttgtagaaaaagtaaaacaataataatagccgctatttggtaaa	4860
Db	4800	ATAATCTATGTATATGTTTAGAAAAAGTTAAACAATTAATATAGCCGCGCTATTTTGGTAAA	4859
Qy	4861	aatccctaataataatcgcgacgatccccgggaattccggggaagcttagatcccatcgag	4920
Db	4860	AATCCCTAATAATCGGACGAGGATCCCGGGAATTCGGGGGAAGCTTAGATCCATGGAG	4919
RESULT	4		
A60112			
LOCUS	A60112	5560 bp	DNA
DEFINITION	Sequence 5 from Patent WO9706267.	PAT	06-MAR-1998
ACCESSION	A60112		
VERSION	A60112.1	GI:3715128	
KEYWORDS			
SOURCE	Transformation vector pTHW142.		
ORGANISM	Transformation vector pTHW142.		
REFERENCE	artificial sequence; vectors.		
AUTHORS	1 (bases 1 to 5560)		
TITLE	De,B.M.		
JOURNAL	Genetic transformation using a PARP inhibitor		
FEATURES	Patent: WO 9706267-A 5 20-FEB-1997;		
source	PLANT GENETIC SYSTEMS NV (BE)		
	Location/Qualifiers		
	1..5560		
	/organism="Transformation vector pTHW142"		
	/db_xref="taxon:126817"		
BASE COUNT	1533 a 1199 c 1244 g 1576 t	8 others	
ORIGIN			
	Query Match 40.7%; Score 2386.8; DB 9; Length 5560;		
	Best Local Similarity 93.5%; Pred. No. 0;		
	Matches 2597; Conservative 0; Mismatches 2; Indels 179; Gaps		
Qy	1	aattacaacggtatatactcgtccagtaactcggcgctgaactcggcgctcagatcgccgtcgagtaacatg	60
Db	1	AATTACAAACGGTATATATCTGTCAGT-----ACTCGGCGCTCGAGTACATG	47
Qy	61	gtcgataagaaaaggaattgttagatgttaattcccatcttgaaagaaataatagtttaa	120
Db	48	GTCGATAAGAAAGGCCAATTTGTATAGATTAAATCCCATCTTGAAAGAANAATATAGTTAA	107
Qy	121	atattttatgataaaataacaaagtcagggtattatagtcacaagcaaaacataaaattatt	180
Db	108	ATATTTATTCATAAAATPAACAAGTCAGGTATTATAGTCCCAAGCAAAACATAAATTTATT	167
Qy	181	gatgcaagatttaattcagaaaaattttcaataactgattatatacagctgggtacattgcg	240
Db	168	GATCAAGATTTAAATTCAGAAATATTTCATAAATACTGATTATATCAGCTGGTACATTTGCCG	227
Qy	241	tagatgaagactgagtcgcatattatgtgtaacataaaattgatgatataagctagctt	300
Db	228	TAGATGAAGAAGACTGATGCCATATTATGTGTATATACATAAATTTGATGATATAGCTAGCTT	287
Qy	301	agctcatcggggatcctagaaacgctgatctcagatctcgtgtacgggagcagaccgac	360
Db	288	AGCTCATCGGGGATCCTAG - ACGCTGTAGATCAGATCTCGGTGACGGGACGACCGGAC	346
Qy	361	ggggcggtacccgagcgtgaagtccagctgccagaaaaaccacgtcatgccaagtcccg	420
Db	347	GGGCGGTACCGGAGGCTGAAGTCCAGCTGCCAGAAACCCACGCTCATGCCAGTTTCCCGT	406
Qy	421	gcttgaagccggcccgccagcatgcccgggggccatatccagagccctcgatcagc	480
Db	407	GCATTGAAGCCGGCCCGCAGCATGCCGGGGGGGCATATCCGAGCGGCTCTGTGCATGC	466
Qy	481	qcacqctcgagtgcttgggcagcccatgacacgacccacqctctcttgaagccctgtgct	540



Qy	2536	tagctaaagctctgttaataactaaagaataatcacacaatctcgaccacgaaaaaac	2595
Db	2687	TAGCTAAAGCTCTGTTAATAACTAAAGAATAATACACAATCTCGACCACGAAAAAACC	2746
Qy	2596	acataataaaatttgaaatt 2613	
Db	2747	ACATAATAAATTTGAATT 2764	
RESULT	5		
AR098311			
LOCUS	AR098311	5560 bp	DNA
DEFINITION	Sequence 5 from patent US 6074876.		PAT
ACCESSION	AR098311		
VERSION	AR098311.1	GI:12807568	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 5560)		
AUTHORS	De Block, M.		
TITLE	Genetic transfection using a PARP inhibitor		
JOURNAL	Patent: US 6074876-A 5 13-JUN-2000;		
FEATURES	Location/Qualifiers		
source	1..5560		
BASE COUNT	1533 a	1199 c	1244 g
ORIGIN		1576 t	8 others
Query Match	40.7%;	Score 2386.8;	DB 9; Length 5560;
Best Local Similarity	93.5%;	Pred. No. 0;	
Matches 2597;	Conservative	0; Mismatches	2; Indels 179; Gaps
Qy	1	aattacaacggtatatactcgtagctactcgccgtcggaactcgccgctcgagtcacatg	60
Db	1	AATTACAAACGGTATATATCTCGCACT-----ACTCGGCGCTCGAGTACATG	47
Qy	61	gtcgataagaaaagcaattgttagatgttaattcccatcttgaaagaaatatagtttaa	120
Db	48	GTGCGATAGAAAGGCCAATTTGTAGATGTTAATTCCTATCTTGAAGAAATATAGTTAA	107
Qy	121	atatatttgataaaataacaagtcagggtatttatagtcacaagcaaaaacataaaattatt	180
Db	108	ATATTTTTCATAAAATAACAAGTCAGGTATTATAGTCCAGCAAAAACATAAAATTTATT	167
Qy	181	gatcaagtttaattcagaataatttcaataactgattatatcaagctgggtacattgcgcg	240
Db	168	GATCAAGTTTAAATTCAGAAATATTTCATAACTGATTATATACAGTCGGTACATTCGCC	227
Qy	241	tagatgaagactgagtcgcgatattatgtgtaatacataaaattgatgatagctagctt	300
Db	228	TAGATGAAGACACTGAGTGCATATTATGTGTAAATACATAAAATTGATGATATAGCTAGCTT	287
Qy	301	agctcatcgggggtactctagaacgcgtgatctcagatctcggtgcgcggcgagaccggac	360
Db	288	AGCTCATCGGGGATCCTAG -ACGCGTGAGATCAGATCTCGGTGACGGCGAGGACCGGAC	346
Qy	361	ggggcggtacgc	420
Db	347	GGGGCGGTACCGGCAGCGCTGAAGTCCAGCTGCCAGAACCCACAGTCATGCCAGTTCCCGT	406
Qy	421	gcttgaagccgc	480
Db	407	GCTTGAAGCCGGCCGGCCGAGCATGCCCGGGGGGCATATCCGAGCGCTCGTGCAATGC	466
Qy	481	gcacgctcggggtcggtgggc	540
Db	467	GCACGCTCGGGTCGTTGGGCAGCCCGCATGACAGCGNACACGCTCTTTGAAGCCCTGTGCTT	526
Qy	541	ccaggggacttcagcaggtgggtgtagacgctggagccagtcgccgcctccgcctgggtggcg	600



[illegible]

QY 1000 agattcaagcgactaggatctacaccgttggtgattttgagtggtgatatgtgtgaggtta 1059  
|||||  
Db 79069 AGATTCAAGTGGACTAGGATCTACACCGTTGGATTTCGAGTGGATGATGTCGAGGTTA 79128  
QY 1060 attttactctggtaacgcccacaaagccctaagagaggtgttgagacccttatcgcttg 1119  
|||||  
Db 79129 ATTTCATTGTTGGTAACGGCCCAAAAGCCCTAAGGAGAGGTGTGTGACACCTTTATCGGCTTG 79188  
QY 1120 aaccgctgggaataatgccacgtggagataatccatgaactcttatcgcttatcatagtt 1179  
Db 79189 AACCCGCTGGAAATAATGCCAGCTGGAAGATAAATCCATGAATCTTATCGTTATCATGAGT 79248  
QY 1180 gaaattgtgtgagtggtgagtggtgctgtcattttacttgctggtggagcttggccct 1239  
|||||  
Db 79249 GAAATTGTGTGATGCTGGAGTGGTCTTCATTTACTTTGCCTGGTGGGACTTTGGCCCT 79308  
QY 1240 ttccctatggggaaatttatattttacttactatagagctttcacaccttttttttaacctt 1299  
Db 79309 TTCCCTATGGGGAATTTATATTTTACTTTACTTATAGAGCTTTTCATACCTTTTTTTTACCTT 79368  
QY 1300 ggaattagtttaataataatggtatgattcatgaataaaaaatgggaatttttgaatttg 1359  
Db 79369 GGATTTAGTTAATATATAATGCTATGATTTCATGAATAAAATGGGAATTTTTCGAATTTG 79428  
QY 1360 tactgctaaatgctaagattagtggaactgtggaatatatatatttttttcaatttaaag 1419  
Db 79429 TACTGCTAAATGCAATAAGATTAGGTGAACCTGTGGAATATATATTTTTCATTTAAAG 79488  
QY 1420 caaaatttgccttttacttagaattataaataagaaaataataataacattcaataaaaa 1479  
Db 79489 CAAAATTTGCCTTTTACTAGAAATTAATAATATAGAAAATATATAACATTCAAATAAAAA 79548  
QY 1480 tgaataaagaacttttcaaaaaacagaaactatgtttaatgtgtaaaattagtcgcacat 1539  
Db 79549 TCAAAATAAGAACATTTCAAAAAACAGAACTATGTTTAATGTGTAAGATTAGTCGCACAT 79608  
QY 1540 caagtcctgtttacatatgttacacaagctcagaagcccaaaagttagcagctcta 1599  
Db 79609 CAAGTCATCTGTACATATNTGTTACACAAGTCTAAGCCCAACAAAGTTAGCACGCTTA 79668  
QY 1600 aataaactaaggtccacgaataattacaaatcataagcccaacaaagtattatgata 1659  
Db 79669 AATAAATAAAGAGTCCACGAAATAATTACAAATCATATAGCCCAACAAAGTTATGATCA 79728  
QY 1660 aaaaaaaacaccccaacaaagctaaacaaagctccaaaaaaactctcaagctcccat 1719  
Db 79729 AAAAAAAAACCCCAACAAAGCTAAACAAGTCCCAAAAAAACTCTCAAGTCTCCAT 79788  
QY 1720 ctctctttatgaacattgaaacctatacacaaacaaagtcagataaactctcttcggac 1779  
Db 79789 CTTCCTTTATGAACATTGAAAACCTATACACAAACAAAGTCAGATAAATCTCTTCTGGGC 79848  
QY 1780 ctgtcttcccaactcctacatacactcctcctatcggattgaattgttttacttgaccttt 1839  
Db 79849 CTGCTTCCCAACCTCCTACATCACTCCGCTATCCGATTGAATGTTTACTTGTACCTTT 79908  
QY 1840 tccgttgcaatgatattgatgatgtttgtgaaactaatagggttaacaatogaagtc 1899  
Db 79909 TCCGTTGCAATGATATGATGATGTTGTTGTGAAAACCTAATAGGGTTAAACAATCGAAGTC 79968  
QY 1900 atggaataggtattggtccaaagattttccgagagctttctcagtagaagaagcccatcaca 1959  
Db 79969 ATGGAATATGGATTTGGTCCAGATTTTCCGAGAGCTTTCTAGTAGAAGCCCATCACA 80028  
QY 1960 gaattactagtataaataaacacaaataggtttcttattatgtgcgaattcaatat 2019  
Db 80029 GAAATTTACTAGTAAATAAATACCAATTAGGTTCTTTATTATGTGTCGCAAAATTCATAT 80088  
QY 2020 aattatagaggatatttcaaatgaaacacgtatgaatgttatagtaaatgtgcaggttaag 2079  
Db 80089 AATTATAGAGGATATTTCAAAATGAAACGTATGAATGTTATTAGTAAATGGTCAGGTAA 80148

QY 2080 acattaaaaaactcactcagtcagatattcacttcaactttaaaaattogattcagtggtgaattgt 2139  
|||||  
Db 80149 ACATTAAAAAATCCTACGTCAGATATTCAACTTTAAAAATTTCGATCAGTGTGGAATTGT 80208  
QY 2140 acaaaaatttgggatcactatataataatgctttcaacacacttggattttttttt 2199  
|||||  
Db 80209 AAAAAAATTTGGGATCTACTATATATATATAAAGCTTTACAACACTTGGATTTTTTTTTT 80268  
QY 2200 gaggctggaatttttaactacatattgttttggccatgcacaaactcattgttttagtg 2259  
Db 80269 GAGCTCGAAATTTTAAATCTACATATTGTTTGGCCATGCACCACTCATTTGTTAGTG 80328  
QY 2260 taatacttgaatttgcataatatatgttcgtgtatatattgtataagaatttctttga 2319  
Db 80329 TAATACTTTGATTTTGTCAAAATATATGTGTCGTATATTTGTATAAAGAAATTTCTTTGA 80388  
QY 2320 ccatacacacacacatatatatatatatatatatatatatatatatcatgcacttttaatt 2379  
Db 80389 CCATATACACACACAT 80448  
QY 2380 gaaaaaataatatatatatatagtcgattttttctcaacacacatatatgttgcgattg 2439  
Db 80449 GAAAAATAATATATATATATATAGTCATTTTCTTCAACACCATATATGTTGCGATTG 80508  
QY 2440 atctgcaaaaactactgtagtaataatgaaaaataataatctattgtgaaattatctcaga 2499  
Db 80509 ATCTGCAAAATACTGCTAGAGTAATGAAATAATATATCTATTGCTGAAATATATCTCAGA 80568  
QY 2500 tgtgaagattttcttaagtaattcttccaaatttttagctaaagtcttggtaaaacta 2559  
Db 80569 TGTTAAGATTTTCTTAAAGTAAATTTCTTCAAAATTTAGTAAAGCTTTGTAATAACTA 80628  
QY 2560 aagaataatacacaaactcgcacacggaagaaaaaacacataataaatttgaatttgcacc 2619  
Db 80629 AAGAATATACAAATCTGCACCAGGAAAAAACACATAATAAATTTGAATTTCTTGACA 80688  
QY 2620 ggc 2622  
|||  
Db 80689 GTG 80691

RESULT 7  
AC004146  
LOCUS  
DEFINITION Arabidopsis thaliana chromosome I BAC F5A8 genomic sequence,  
complete sequence.  
AC004146  
AC004146.1 GI:3540210  
VERSION  
KEYWORDS HTG.  
SOURCE  
ORGANISM Arabidopsis thaliana  
thale cress.  
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 101933)  
Fedorispiet,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R.,  
Au,M., Araujo,R., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y.,  
Oji,O., Osborne,B.I., Shinn,P., Sun,H., Toriumi,M.,  
Vyotskaia,V.S., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.  
Unpublished  
2 (bases 1 to 101933)  
Fedorispiet,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R.,  
Au,M., Araujo,R., Brendel,V., Buehler,E., Dewar,K., Feng,J.,  
Kim,C., Li,Y., Oji,O., Osborne,B.I., Shinn,P., Sun,H., Toriumi,M.,  
Vyotskaia,V., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.  
Direct Submission  
Submitted (31-JUL-1997) Biochemistry, Stanford University/PDA  
Sequencing and Technology Center, 855 California Avenue, Palo Alto,  
CA 94304, USA  
3 (bases 1 to 101933)  
Fedorispiet,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R.,  
Au,M., Araujo,R., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y.,  
Oji,O., Osborne,B.I., Shinn,P., Sun,H., Toriumi,M., Vyotskaia,V.,

Yu,G., Ecker,J., Theologis,A. and Davis,R.W.

Direct Submission  
Submitted (04-SEP-1998) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA

REFERENCE 4 (bases 1 to 101933)

Federspiel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R.,  
Au,M., Araujo,R., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y.,  
Oji,O., Osborne,B.I., Shinn,P., Sun,H., Toriumi,M., Vyotskaia,V.,  
Yu,G., Ecker,J., Theologis,A. and Davis,R.W.

Direct Submission  
Submitted (30-JAN-1999) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA

On Sep 4, 1998 this sequence version replaced gi:2934888.  
Bases 90722-101933 of clone F5A8 overlap with bases 114738-103537  
of 'IGF' BAC clone F1M21.

FEATURES

source

1..101933

/organism="Arabidopsis thaliana"

/cultivar="Columbia"

/db\_xref="taxon:3702"

/chromosome="I"

/clone="IGF' F5A8"

complement(2056..2841)

/gene="F5A8.1"

/note="ribulose biphosphate carboxylase, small subunit;  
ats1a"

complement(join(2057..2293,2430..2564,2671..2841))

/gene="F5A8.1"

/EC\_number="4.1.1.39"

/note="Location of EST gb|AA042768. atsl1a gene  
emb1|X13611"

/codon\_start=1

/product="ribulose biphosphate carboxylase, small  
subunit"

/protein\_id="AADI0655.1"

/db\_xref="GI:4204274"

/translation="MASSMLSSATMVASPAQATMVAFNGLKSSAAPPATRKANNIDIT  
SIFNGRVMCMWPPIGKKFETLSYLPDLTSELAKEDVDYLIRNKWIPCVFEFLE  
HGFVYRHGNSPGYDGRYTMKPLFGCTDSQAQVLKEVECKEYKPNAFIRIIGFD  
NTRQVQISIFIAKPPSFTG"

7393..8251

/gene="F5A8.2"

join(7393..7582,7740..8251)

/gene="F5A8.2"

/note="Hypothetical protein"

/codon\_start=1

/protein\_id="AADI0658.1"

/db\_xref="GI:4204277"

/translation="MRMSCNGCRVLRKGCSENCIRPLOWIKSAESQANATVFLAKF  
YGRAGLMLNTGPDHLRPAIFRSLLEYACGRIVNPIYGSVGLWSGNHLCQAAVEA  
VMRGSPTPTACDAAVTGOAPPENNKICDIRHVSRDENVKRRSRGACKERNVRSL  
HESSLHESVPSSEETTEPQWIGLILGLPLELARGNHVVVPMKRLKRCGTSE  
DEDTCRTELGLVCE"

complement(15428..18947)

/gene="F5A8.3"

complement(join(15428..15850,17264..17636,17980..18439,  
18665..18947))

/gene="F5A8.3"

/note="putative Cytochrome P450 protein; Similar to  
Catharanthus roseus Cytochrome P450 gi|404690 and others.  
Contains Cytochrome p450 motif"

/codon\_start=1

/protein\_id="AADI0659.1"

/db\_xref="GI:4204278"

/translation="MMVTLVLKYLIVMTLILRVLYDSICCYELTPRRIKKFMERQG  
ITGPKPLRGTNIIDISKMLSHSANDCSSIHINIVPRLLPHYVSWSKQYKGRIMNN  
GTEPRCLTETEMIKHTKHPVTGKSWLQOQQTGKFIGRGLMAGAAWHHORHMA  
APATDRLLGAKYKHWECTKMAERLKEVEGEIEMRRLTADIISFTFGSSC  
DKGKELSLTLVLRCAQATRLCLFCPSRFLPSKYNREIKSLKTEVERLLMEIDSR  
KDSVEIGRSSYGGDLLGLLQNDKNNLNQVIMDECKTFTTGHETYSLLLTWT  
LMLLAHNPTQDQNVDRVQVCGODGVPSEVLEQSLSLTLNKNVINESLRLEYPATLLPR

MAFEDIKGLDIIIPKGLSIWIPVLAIHHSNELWGEDANENPRTTTRSFASSRHFWP  
FAAGPRNCIGQTAMMEAKIILAMLVSKFSFAISENYRHAPIVVLTIKPKYGVQLVK  
PLDI"

complement(23318..49118)

/gene="F5A8.4"

complement(join(23319..23351,23434..23565,23671..23949,  
24045..26054,26197..26352,26554..26688,28076..27054,  
27106..27384,27459..27617,27707..27928,287018..28251,  
28349..28551,28655..28749,28844..29001,29141..29297,  
29386..29584,29834..29894,29992..30104,30235..30310,  
30517..30681,30769..30899,30985..31084,31150..31269,  
31351..31584,31660..31761,31889..32071,32153..32332,  
32421..32570,32676..32789,32881..33035,33145..33268,  
33405..33494,33957..34094,34143..34228,34305..34437,  
34601..34849,34927..35006,35322..35387,35476..35653,  
35739..37247,37375..38246,38336..38543,38649..38930,  
39082..39171,39254..39354,39437..39535,39649..39790,  
40066..40139,40861..41011,41411..41465,41548..41741,  
41894..42199,42289..42393,43162..44931,45058..45180,  
45259..45432,45649..45768,45852..45921,46007..46106,  
46179..46301,46660..46738,46894..47013,47108..47217,  
47445..47614,48094..48266,48379..48463,48588..48913,  
49014..49118))

/gene="F5A8.4"

/note="Hypothetical protein; Similar to Saccharomyces  
cerevisiae probable membrane protein YLR106c, gi|1360494"

/codon\_start=1

/protein\_id="AADI0657.1"

/db\_xref="GI:4204276"

/translation="MAIDGSFNKLALETFVSRCPKVAAPFCFTSILSKGGEVDNEE  
VIHALGDAFLHPFTVPLVHCFPIIRNVDRVVLGLRLVDLKLSSIDYSDDVSSVLD  
NAMTEGISVIDFYVRQGRQLHECECLAFRAHLHFNTSLIGLSILNIEKAPPERYI  
LVKDIVSESRMEATDAYLLCLRVSRYFLVIRPEVFSKLDMSCLVDSKRLSECPROQ  
RHLFLEKYRDVMDVNSFEPEIHSRVKKSPEMSLAVSQRPVLLYGPSSGSKSALI  
RKLADSESNHVFIHMDOLDGKTLVGTCTDQDQGERFWQSGIQTAIMNGFWVLE  
DIDKAPSDVPLVLSLLGSCSFLTSGEEIRIAETFLQFSTISTPECSVSHIRDAQN  
SLDKPVRVVPDPRESLQILGARYPLNGPVAEKLJETFETINSALRQPFSSSTTE  
NSATFSPSPRESLRDLKWCERVHGPSYDGHAVYQEAADIFSANMSVKNRAVSEI  
VASTNVAVPESQDKPQIHSRDRSFVETSTRLEKIIARSVEYNPVLVGETGTG  
KTTLVQNLAHWIGOKLTVLNSQSDIVDLGFGKPIDPLKMTWYNEFNEALRDLC  
IKDDSKLMKLQDNFRACKWHTFLTGLADI1KGLEGRITERMEGKIGARSRSRKRK  
KPEBELNCACLRKYNKIROQIIHSGGNVTFVEGAFVTLREGHWLLDEVNLAPPE  
ILGLVIGLEGRSGLAERGDWNGIPRHLNRLFACMNPATDAGKRDLPFSFRSEI  
TEYADDDICDDLEIFVYRLEGGSGSLVANIVWFYKEAKRLEESLQDSQANOKP  
QYSLRSLRAYEAIKABIGAFQKALYDFSMFFLSLLDASSAKIVBPIIKRISGEN  
IRSQPORYLCELKGGSDKFGYSVKYKTSVDHLNHLAHAIFIKRYPVLLQPTSSGK  
TSLVKYLAASSENKFRVNNHEQTDIOEYLGYSMTDSSGKLVFHEGALVAVRGHWI  
VLDELNLAPSDVLEALNRLDLDNRELFVPELSETLSAHPNMFATQNPPLYGGRKI  
LSRAFRNNRVEIHDPEDEISELITTKCSIANSHKSMVEMKLDQNRQSSKAF  
GKHGYITPRDLFRWAYRFTYDGTSHIELAREGYIYLAERLDDETKEYVQVEVERHF  
RVSLAKDLYNNGFFPVDRSKLITEYENOVQLELSQALTPFGODIVICGDISRAEV  
SIKSVEVALEKYKNGSVTGAATQDDVDFLEKIRNNVMVLYQKRAIVFQDGPGLVEA  
MAGNIVLVDEISLADDSVLRMNSVLETDRLKSLAEKGPVLEVAHEDFFVLATM  
NPGDYGKKELESPALNRFTIWWPTTDEELRSIAFSGLSLLESNVVDPIINFWE  
WFNLHGTRTLTVLRLDLSVAFNAMESLGPAYAILHGAFVLVLDGLSLGTGFSGRD  
GODLREKCFAPLLQOELFASDTPLSELSRMELYGWDGSAICKEKSVRHEGMFIGD  
PFFTSKGENPEIIGFEPFAPTHRNRLVRAMQLSKPILLEGSPGVGKSTLILAG  
KYSKGVVYRINLSQTDMDLGLSDLPVEDSEDMKFWSGDGLLQGLNALIDHRAOVF  
IPELGCTFECPTFRVACQNPSTGGGKGLPKSFLNRFTKVYVDELVEDDYLFICR  
SLYFSVPSLLSKLIALNRQHDGTMLYRKFGHDSGPWFNRLDVRVIRSCQPMQERAIH  
LEVFSFLNVLYIQMRMTDRKEVRIYKAIFDKPTPSINPRVONPAYLVVGTAAI  
KRNLSQNIASEQLKLLPEIRQNEAVAHCVONKWCILVGPSSGKTSVIRILIAOLF  
GYPINELNLSATSDSLGCGFQYNAFNRFLVMTRVLEHVDVYNSLLSQSEALF  
SNRSLGSRWLSYLNKIDSSIVENPFLFNDSETLLEEVDELEVLYKKGVLPSW  
SKKYLEQISKTLIQLOTHEKQSTKEFWTGMKLIKAEKGEWVLYKNANLCNPVLD  
INSLVPCGSLITINECGLVNPEPTVVPHPNRLFLSNPKFGEVSRAMNRGVEVFM  
MGPHQWNEDSGNCEELVLRGVFSLGSPGYKLVTSMAKHAHVHAWLNGSGFVRI  
TYLEQWVHLFQLLMMNGNLLMSLQSWEHYLSLSLGTVDKGEVDFVRETYLSDV  
ELSELDSPRDELAAGEPRITYSIDSWTLKKVLPKALIGSSCAPADANENDLAKMLL  
SDNKYSRMDRELAAGEPRITYSIDSWTLKKVLPKALIGSSCAPADANENDLAKMLL  
FAANWTTEQATEEDIQLYLAWFSWFGSRLOQHCPLLCFLNTLKVFEFHPVNHHSRC  
RKNKFLCRLDQDAVPIPMLSKSLIDVANDQSKAPYSKLSFESLNSVGVLRHSFQW  
LVESNDNHTDVTSTTRFLDLSURVLEKKTLCEITVGAPSFVLTLQLYTEVIDNHSFFWSG



LVSSDSEYLLSEFSLIKSIKKMHISFFPEQVVLSESKNNINVLHGHPKSKMLWAY  
 GCHPSLPVASEDFHKKQEFLOLCSTVMPKSEDEHNDHUTKAIPSPGPELCLLAE  
 GLISSYLADEDDVDYVAALDEIYOTFLERLKLKKRLKMDKMGFSEIDONTENITAS  
 CVCPCPELVITVSGSFSSMWKTCFTASSESCSDVELLAALHQLHLLVAPTEHQDLVDLR  
 KLLAPALYSLSRPPOTLVAHQKLLWADAHASELGVDTKIAGLEIWHWSVL  
 WKNQOIGLMAHLVTSFFHLMFTGPTFLRHQWLSDSVTFYADACSNYSRCSIIY  
 THQSFSENEFVAKSVFHALEKKONKMDGIONLISLIGSSHNKLSKVTHSFVGPPLA  
 LNIYRSGEYFSLGLLYSGNDESEHTLSKLTETKRLORKVIFRSDPKPKYQDLRA  
 LDEPAGFTLRPISVNDIEVDNMVQVFNQOETAKLSTFRDMSSDYSEVVDITQPI  
 QVSVYKMLKLSVSGALLGLKLNRFIDIMVDSMETIYALMRPFRDSSIASTTYTE  
 CLPPLHLSHGANSRAKSLIGLDVGLLHKLISVSSAEDSRKASELOKLKYLXNLHARVL  
 QFVANTGLLDEASELQDKIYELARLWEMKFOAKTKADNLPGLYFRSRDRKIDSV  
 MEVDLSALGTFPFPNSESSEWQYLAADDOTKNVDMTHIDQDEENLEDWDLIOHLDLS  
 IYTHNLFELFCDSERTIYISIQSRECFITDSRLRDSFTDSYELGVSMIKGLRGLFVS  
 SLDAKLVPEHLRLCLNKNKNTSNYOASKINPYKOLDGPELGMKVFETPLQOORN  
 SLQEREDHPGLQKLSVLOMLLAIPSSTPLAKALGLOPLCKVHKLOEGBGKLPIIS  
 DLLPEPIISLASSMOKVFERPWTLLDVOQOYELNARKFCFRMLKWFQNMKYSPPHK  
 VSFTLKLKGLQHYLLKSSQLIAPPLFQLTSNLTVDFSRYFSDSHKRWEMVCYNLF  
 GFYIQFVPVHEQDLNKNVETELKVLKLRWERPDNLYNETTKRTRQKVKKLIQ  
 KFTDMLRPLVNLKVPDUTKERAQFLPLDPLDMDGASDMRIEVLVSALDAEQLDRSS  
 WYVVMNKLKESVRFQHEMYKTLLMGAHQYSSPYQGDWKNLWSTVARIGETIAG  
 CSDLWRNSDRDRAKRALFELLKLESSGLQKHKFENIEMSNHFKGLLYXOPAYDPKHL  
 LLLTTHSNHPSMGVEDQNKENSLEVMRYANEFYFKSLASVQLMLNIDRKHSDVTAE

Query Match 29.68; Score 1736.6; DB 12; Length 101933;  
 Best Local Similarity 99.8%; Pred. No. 4.1e-248;  
 Matches 1739; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy	880	ccattgttctcttactcttctgtgactgaggttggcttagtctggttctggtcatctat	939
Db	2838	CCATTGTCTTCTTACTCTTGTGACGAGGTTGGCTAGTCTTGGTTCATCTAT	2897
Qy	940	atataatgatatacaacaatgagaacagcttggagtgatcgagggttctaggatacatg	999
Db	2898	ATATAATGATAACAAACATGAGAACAGCTTTGGAGTGATCGGAGGCTTAGGATACATG	2957
Qy	1000	agattcaagtgagctagagctcacacgttggatttggatgtgagatgtgtaggtta	1059
Db	2958	AGATTCAAGTGGACTAGGATCTACACCGTTGGATTTTGAGTGTGGATGTGTGAGGTTA	3017
Qy	1060	attttacttggtaagggcccaaaaggcctaaggagaggtgttgagacccttatcggttg	1119
Db	3018	ATTTTACTTGGTAACGCCACAAAGGCTTAAGGAGAGGTGTGTGAGACCCCTATCGGCTTG	3077
Qy	1120	aaccgctggaataatgccagtggaagataattccatgaatcttctatgatt	1179
Db	3078	AACCGCTGGAATAATGCCAGTGGAGATAAATCCATGAATCTTATCGTTATCTATGAGT	3137
Qy	1180	gaaattgtgtgattgtggagtggtgtgctgtcattttacttgcctgctggtgacctggccct	1239
Db	3138	GAATTTGTGTGATGGTGGAGTGGCTGTGCTCATTTTACTTTGCTTGGTGACCTTGGCCCT	3197
Qy	1240	ttcctttaggggaatttataattttacttactatagagctttcatcaccttttttttaccctt	1299
Db	3198	TTCCCTTATGGGGAATTTATATTTTACTTACTATAGAGCTTTCATACCTTTTTTTACCCTT	3257
Qy	1300	ggattttagttaataataatgattgattcattcgaataaaataatgggaaatttttgaatttg	1359
Db	3258	GGATTAGTTAAATATAATAATGGTATGATTTCATGAATAAAATGGAAATTTTGAATTTG	3317
Qy	1360	tactgctaaatgcataagattagtgaaactgtggaataataatttttttccattttaaag	1419
Db	3318	TACTGCTAAATGCATAAGATTAGGTGAAACTGTGGAATATATATTTTTCATTTAAAAG	3377
Qy	1420	caaaatttgccttttactagaataataataatagaataataataacattcaataaaaa	1479
Db	3378	CAAAATTTGCCTTTTACTAGAAATTAATAATAGAAAAATATATAACATTCATAATAAAAA	3437
Qy	1480	tgaataataagaactttcaaaaaacagaactatgttttaattgttaaagattagtcgaact	1539
Db	3438	TGAAATAAAGAACTTTCAAAAACAGAACTATGCTTTAAATGTGTAAAGATTAGTCGCACAT	3497

Qy	1540	caagtcattctgttacaataatattgttacaacaagtcataagcccccaaaaagttagcacgtcta	1599
Db	3498	CAAGTCATTCTGTTTACAATATGTTTACACAAGTCATATAAGCCCAACAAAGTTTAGCAGCTCTA	3557
Qy	1600	aataaactaaagagtcacacgaaaattattacaaaatcataagcccccaaaaagttagtatca	1659
Db	3558	AATAAACTAAAGAGTCACAGAAAATATTACAAATCATTAAGCCCAACAAAGTTATTGATCA	3617
Qy	1660	aaaaaataaacgcccccaacaagaagctaaacaagctccaaaaaaacttctcaagctccat	1719
Db	3618	AAAAAATAAACGCCCCAACAAAGCTAAACAAGTCCAAAAAATACTTCTCAAGTCTCCAT	3677
Qy	1720	cttccttatgaacattgaaactatacacaaaacaagaagtcagataaaattctctttctgggc	1779
Db	3678	CTTCCTTTATGAACATTGAAAACATATACAAAAACAAGTCAGATAAAATCTCTTTCTGGGC	3737
Qy	1780	ctgtcttcccacacctctacatcactctccctatcgattcgatttgaattttttactgttacctt	1839
Db	3738	CTGTCTTCCCACAACTCCTACATCACTTCCCTATCGGATTGAATGTTTACTTGTACCTTT	3797
Qy	1840	tccttgcgaatgatattgatagtagtattgttgaaaaactaatagggtttaacaatcgaaagtc	1899
Db	3798	TCCGTTGCAATGATATTGATAGTATGTTTGTGAAAACTAATAGGGTTTAAACAATCGAAGTC	3857
Qy	1900	atggaatatgagattggtgccaaagattttccgagagctttcttagtagaaagccccatcaaca	1959
Db	3858	ATGGAATATGGATTGGTCCAGATTTTCCGAGAGCTTCTTAGTAGAAAGCCCATCACCA	3917
Qy	1960	gaaattactagtaaaaaataataaccacaattagggtttcttatttatgtgccaaaattccaatat	2019
Db	3918	GAAATTTTACTAGTAAAAATAAATCACCAAATAGGTTTCTTATTATGTGCCAAATTCATAT	3977
Qy	2020	aattatagagatatattcaaatgaaaacgtagaatgttatttagtaaaatggcaggaag	2079
Db	3978	AATTATAGAGGATATTTCAAATGAAAAACGATGAATGTTTATTAGTAAATGGTCAGGTAAG	4037
Qy	2080	acattaaaaaaactcctcagtcagatatccaaacttcaaaattcgatcgagtggtgaaattgt	2139
Db	4038	ACATTTAAAAAATCCTACGTCAGATATTCAACITTTAAAAATTCGATCAGTGTGGAATGT	4097
Qy	2140	acaaaatttgggatctactatataataataaagcttttcaaacacacttggatttttttttg	2199
Db	4098	ACAAAAATTTGGGATCTACTATATATATATATAATTAATGCTTTACAACACTTGGATTTT	4157
Qy	2200	gaggtggaatttttaactacatatattgttttggccatgcacacactcattgtttagtg	2259
Db	4158	GAGGCTGGAATTTTAACTCTACATATTGTTTGGCCATGCACCACTCATTTGTTTAGTG	4217
Qy	2260	taatactttgattttgtcacaataatagtgttcgtgtatatatttgtaagaatttctttga	2319
Db	4218	TAATACTTTGATTTTGTCAAAATATATGTTGCTGTGTATATTTGTATAAGAATTTCTTTGA	4277
Qy	2320	ccatatcacacacacatatatatatatatatatatatatatatatatatatcatgcacttttaatt	2379
Db	4278	CCATATACACACACACATATATATATATATATATATATATATATATATATATCATGCAC	4337
Qy	2380	gaaaaataatatatatatatagtcgatttttttcaaacacacacacacacacacacacacacac	2439
Db	4338	GAAAAATAATAATATATATATATATAGTGCATTTTCTTAACAACACATATATGTTGCAATG	4397
Qy	2440	atctgcacaaaatactgtagagtaagtaaaataataatctattgtcgaaatttatctcaga	2499
Db	4398	ATCTGCAAAAATACTGCTAGAGTAATGAAAAATATATATCTATTTCGTGAAATTTATCTCAGA	4457
Qy	2500	tgtaagattttcttaaaagtaaatcttttcaaaatttttagctaaagctcttgtaataacta	2559
Db	4458	TGTTAAGATTTTCTTAAAGTAAATTTCTTTCAAAATTTAGCTAAAAGCTCTTGTAAATAACTA	4517
Qy	2560	aagaataatacaacaactctcgacacgaaagaaaaaacacacacacacacacacacacacacac	2619
Db	4518	AAGAATAATACAAATCTCGACACGAAAAAACAACACATAATAAATTTGAATTTCTTGACA	4577
Qy	2620	gcg 2622	





QY	4141	tcctacgagagatagattgaaagaagtgcagggaagcggttaactcggagaacatacaacaat	4200
Db	754	TCATTCGAGAGATAGATTGAAGAAAGTCGAGGAACGGTTAACTCGAACATAACACAAT	695
QY	4201	gtctaataatgcattcgcataaccaaagaagtgtattactctctccggtccacaataag	4260
Db	694	GTCTAAATTAAATGCAATTCGCTAACCAAAAAGTGATTACTCTCTCGGTCACCAATAAG	635
QY	4261	ttatttttggccctttttttatggtccaaaaaagaatgagtttttttagatttccaaaaatg	4320
Db	634	TTATTTTGTGCCCTTTTATGTTATGTCCTCAAAATAAGTGAGTTTTCATATTCAAAATG	575
QY	4321	atttaattatttttttactacaagtcccttggagtaaaatggttggagtagtatgttaga	4380
Db	574	ATTTAATATTTTTTACTACAGTGCCTTGGAGTAAATGCTGTGGAGTATGTGTAGA	515
QY	4381	aatgtttatgtgaagaataagtaaaaggttaatatgatcaattcattcgtatttaattgtt	4440
Db	514	AATGTTTATGTGAAGAAATAGTAAGGTTAATATGATCAATTTTCATTTGCTATTTAATGTT	455
QY	4441	aaaatgtgaatttcttaattctgtgtgaaacaaacccaaaaatcacattatttggaccgga	4500
Db	454	AAAAATGTGAATTTCTTAATCTGTGTGAACCAACCAAAAAAATCACTTATTGTGACCGGA	395
QY	4501	gaagatataaataatatatttggagcgactaaaaataaaacttttctcatatatacga	4560
Db	394	GAAAGTATATAAATATATATTTTGAAGCGGACTAAAAATAAACTTTTCTCATATTATACGA	335
QY	4561	acctaaaaacagcatatggttagttcttagggaaatcctaaatcactaaataataaaagaa	4620
Db	334	ACCTAAAAACAGCATATGTTAGGGAATCTAAATCACTAAAAATTAATAAAGAA	275
QY	4621	gcacaagtagtaacatcatatgatattcacaccgtcaaacacgaaatctgtaaatatattta	4680
Db	274	GCAACAAGTATCAATACATATGATTACACCGTCAACACGAAATTCGTAAATATTTTAAT	215
QY	4681	ataataaagaattaatccaaatagctccaccctataactaaactaaaaataaacagc	4740
Db	214	ATAATAAAGAATTAATCCAAATAGCTCCACCCTTAACCTTAAACTTAAAAATAACACG	155
QY	4741	gaatgatattatcatgataatttatattataatgaatggtataatcatgataatcaatgt	4800
Db	154	GAATGTATATTATATGCATAAATTTATATATTAAATGTGTATATCATGTATATCAATGT	95
QY	4801	ataatctatgtatggttagaaaaagtaaacaaattaataatagccggctatttgtgtaa	4860
Db	94	ATAATCTATGTATGTTAGAAAAAGTAAACAATTAATATAGCGGCTATTTGTGTAAA	35
QY	4861	aatccctaataataatcg 4877	
Db	34	AATCCCTAATATAATCG 18	
RESULT 10			
A10061/c			
LOCUS	A10061	3266 bp	DNA
DEFINITION	N.tabacum (pta2953), TA29 gene.		
ACCESSION	A10061		
VERSION	A10061.1		
KEYWORDS	Common tobacco.		
SOURCE	Nicotiana tabacum		
ORGANISM	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.		
REFERENCE	1 (bases 1 to 3266)		
AUTHORS	Martani,C., Leemans,J., De Greef,W. and De Beuckeleer,M.		
TITLE	Plants with modified stamen cells		
JOURNAL	Patent: EP 0344029-A 9 29-NOV-1989;		
PLANT	GENETIC SYSTEMS, N.V		
FEATURES	Location/Qualifiers		
source	1..3266		
	/organism="Nicotiana tabacum"		

gene	/db_xref="taxon:4097"			
CDS	1527..2492			
	/gene="TA29"			
	1527..2492			
	/gene="TA29"			
	/codon_start=1			
	/protein_id="CAA00875.1"			
	/db_xref="GI:490022"			
	/db_xref="SWISS-PROT:P24804"			
	/translation="MVAPKWFISFMILLSLAICSGQPVTSDAIKAKEADHNLKAHT LSNIDAKFGGGFGIGGAGGGGGSDPTNPGYNGPSCIHGCTVPGFGFL PXPVFCVPVSPCGVCVPADIPGTMETSKITIGSARLYRCKPQNMCDSKDQNE LLLLHFVPMQDKHDKQEHRLRYGRRIGLITVGVGGFGIGFCANGGGGGGGGSDA PGSNDGCDPFGCGPCGYACPANPNPSSGITEFHSLSRFDGPRYCRPDMCESEDC NEULLHFVSPMQHKHNRHDHIVVERDEEAAHQSHQHKHDEDIIN"			
BASE COUNT	1016 a	581 c	623 g	1046 t
ORIGIN				
	Query Match	25.8%;	Score 1512.2;	DB 9; Length 3266;
	Best Local Similarity	99.8%;	Pred. No. 1e-214;	
	Matches 1514;	Conservative 0;	Mismatches 3;	Indels 0; Gaps 0;
QY	3361	ggtaccatggtagctaatttctttaagtaaaaaactttgatttgattgattgattgttact	3420	
Db	1534	GCTACCACTTTAGCTAATTTCTTTAAGTAAATAAACTTTGATTTGAGTGATGATGTTGACT	1475	
QY	3421	gttacacttcacacacaaaggccatatatagagcacaaagacatacacacaacttgcacaaa	3480	
Db	1474	GTTACACTTCACCACCAAGGCGCATATATAGAGCACACAGACATACACACAACACTTGCANAA	1415	
QY	3481	ctaacctttgttgagcatcttcagaggaataatggggagtagcaggcctaactcgagggtaac	3540	
Db	1414	CTAACTTTTGTGGAGCATTTTCGAGGAAATTCGGGAGTAGCAGGCTAATCTGAGGGTAAC	1355	
QY	3541	attaagtttcacgtatttaatttgttcacaaatggacttagtgtgagggaaaaagtagacca	3600	
Db	1354	ATTAAAGTTTCATGTATTAATTTGTTGCAACAATGAGACTTAGTGTGAGGAAAAAGTAGCCA	1295	
QY	3601	aaattttgtctccacctgatttcagttatggaaattacattatgaagctgtgcagagaaa	3660	
Db	1294	AAATTTTGTCTACCCCTGATTTTCAGTTATGGAATTTACATTTATGAAGCTGTGCTAGAGAA	1235	
QY	3661	gagttttattctagtcagccaccaccttatgcagctgtcttttagcttgattcctcaaaa	3720	
Db	1234	GATGTTTATTCTAGTCCAGCCACCACCTTATGCAAGTCTGCTTTAGCTTGATTCANAA	1175	
QY	3721	actgatttaattacattgtctaaatgtgcatacttcagagcctatgtcgccttaattcgag	3780	
Db	1174	ACTGATTTAATTTTACATTTGCTAAATGTGCATTCGAGCCTATGTGCTTTAATTCGAG	1115	
QY	3781	taggatgtatatattagtcataaaaaatcatgtttgaaatcatcttccataaaagtgcacaa	3840	
Db	1114	TAGGATGTATATATTAGTACATAAAAAAATCATGTTTGAATCATCTTTTCATATAAAGTGACAA	1055	
QY	3841	gtcaattgtccctcttctgttggcactatatattcaactctgttaactgttaagcaaatatccagtta	3900	
Db	1054	GTCAATTGTCCCTTCTTGTGTTGGCACTATATTCATCTGTTAATGCAAAATTTATCCAGTTA	995	
QY	3901	tacttagctagatataccaatttttgaataaaaaatagctctcttgattagtaaacccgagatagt	3960	
Db	994	TACTTAGCTAGATATCCAAATTTTGAATAAAAATAGCTCTTTGATTTAGTAAACCGGATAGTG	935	
QY	3961	acaaagtccatatccatcaaaactctgtgctcgtggtctaaagtctctgacgcacatgggg	4020	
Db	934	ACAAAGTCACATATCCATCAAACTTCTGTGCTCGTGGCTAAGTCTTGATCGACATCGGG	875	
QY	4021	ttaaaatttaaaattgggacacataaatagcctatttgtgcacaaatctccccatcgacaaatg	4080	
Db	874	TAAAAATTTAAATTTGGACACATAAATAGCCTATTTGTGCAAAATCTCCCCATCGAAAATG	815	
QY	4081	acagattgttacatggaaaaacaaaagtcctctctgatagaagtcgcaaatgtatcaccaattt	4140	

Db 814 ACAGATTGTTACATGAAACAAAAAGTCTCTGTAGATGCGCAAGATATCACAAATTT 755  
Qy 4141 tctatcgcagagatagatgaaagaagtcacgggaagcggttaactggaacataacacaat 4200  
Db 754 TCTATCGAGAGATAGATTGAAGAAGTCGACGGGAAGCGGTAACTGGAACATAACACAAT 695  
Qy 4201 gtctaaataattgcatcgcgttaaccacaaagtgattactctctccggtccacaataag 4260  
Db 694 GTCTAAATTAATTGCTATCGCTAACCAAAAAAGTGATTACTCTCTCGGTCACCAATAAG 635  
Qy 4261 ttatttttggcccttttttttatgggtccaaataaagtgagtggttttttagatttcaaaatg 4320  
Db 634 TTATTTTGGCCCTTTTATGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 575  
Qy 4321 atttaattatttttttactacagtgcccttgagtgagtaaatggttgagtgagtgagtgag 4380  
Db 574 ATTTAAATTAATTTTACTACAGTGCCTTGGAGTAATGGTGAGTGGAGTGGAGTGGAGT 515  
Qy 4381 aatgttttatggaagaataagtaaggttaatatgatcatatttctcatttcttatttattgtt 4440  
Db 514 AATGTTTATGTGAAGAATAAGTAAAGGTTAATATGATCAATTTTCTCATATATACGA 455  
Qy 4441 aaatgtggaattcttcttaattctgtgaaacacacacacacacacacacacacacacac 4500  
Db 454 AAATGTGAATTTCTTAATCTGTGTGTAACCAACCAACCAACCAACCAACCAACCAAC 395  
Qy 4501 gaaagtataaaatatatttggagcgactaaacacacacacacacacacacacacacacac 4560  
Db 394 GAAAGTATATAATATATATTTGGAAGCGACTAAAAATAAATTTTCTCATATATACGA 335  
Qy 4561 acctaaac 4620  
Db 334 ACCTAAACACACATATGTTAGTATTTCTAGGGAATCTAAATCACTAAATTAATAAAGAA 275  
Qy 4621 gcaacaagtatcaatacatatgatttacacgcgtcaaacacacacacacacacacacacac 4680  
Db 274 GCAACAGTATCAATACATATGATTACACCGTCAACACAGAAATTCGTAATATTTAAT 215  
Qy 4681 ataataaagaataaataccaaataagcctccacacacacacacacacacacacacacacac 4740  
Db 214 ATATAAAGAATTAATCCCAATAGCTCCACACCTAACTAACTAACTAACTAACTAACTAA 155  
Qy 4741 gaatgtatattatgataataatttatataatgataatgataatgataatgataatgataat 4800  
Db 154 GAATGTATATTATATGCATAATTTATATATTAATTAATGATATATATATATATATATAT 95  
Qy 4801 ataataatgataatgataatgataatgataatgataatgataatgataatgataatgata 4860  
Db 94 ATAATCTATGTATGTTAGTAAAGTAAACAAATTAATATAGCCGCTATTTGTGTAA 35  
Qy 4861 aatccctaaataataatcg 4877  
Db 34 AATCCCTAATATAATCG 18

NTTA29 6254 bp DNA PLN 09-MAY-1995  
Tobacco anther-specific gene TA-29 and stem-specific gene TSJT1.  
X52283  
X52283.1 GI:20035  
tissue specific gene.  
common tobacco.  
Nicotiana tabacum  
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids  
I; Solanales; Solanaceae; Nicotiana.  
1 (bases 1 to 6254)  
Mariani, T.  
Direct Submission  
Submitted (23-MAR-1990) Mariani T., Plant Genetic Systems, J  
plateaustraat 22, B 9000 Gent, Belgium  
2 (bases 1 to 6254)

RESULT 11  
NTTA29/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE

AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
COMMENT  
FEATURES  
source  
Location/Qualifiers  
1. 6254  
/organism="Nicotiana tabacum"  
/cultivar="Samsun"  
/db\_xref="taxon:4097"  
/clone\_lib="lambda Charon32"  
1446..1452  
/note="TSJT1"  
1477..2714  
/gene="TA29"  
1477..(2617,2714)  
/gene="TA29"  
1527..2492  
/note="anther-specific"  
/gene="TA29"  
/note="anther-specific"  
/codon\_start=1  
/protein\_id="CAA36524.1"  
/db\_xref="GI:20036"  
/db\_xref="SWISS-PROT:P24804"  
/translation="MVAPKWFISPMILLSLAICSGQPVTSDAIKAKEADHDLKAHT  
PKPVDFGVYSPGGYGVGADPTGGMTESKIGTSQSARLYRCKPGNMGDSKDCNE  
LLLFVPMQDKHDKHRLYGGRRIGLTVGGVGGFCIGCGAMGGGGGGGGGSDA  
PCNSNDGCDPGFCGPGCYACAPNPNPSSGITEPHISGLSRFDGPKRCRPMCESED  
NELLLHFVSPMQHKHNRHDIHVERDEEAHHOSKHQKDEDIIN"  
1527..2489  
/gene="TA29"  
/product="anther-specific"  
complement(<3022..>4678)  
/gene="TSJT1"  
/note="stem specific, weakly expressed in other organs"  
complement(<3022..>3075)  
/gene="TSJT1"  
/number=3  
complement(join(<3022..3075,4047..4287,4473..>4678))  
/gene="TSJT1"  
/note="stem specific, weakly expressed in other organs"  
complement(3022..4678)  
/gene="TSJT1"  
complement(join(3022..3024,4047..4287,4473..4678))  
/gene="TSJT1"  
/note="stem specific, weakly expressed in other organs"  
/codon\_start=1  
/protein\_id="CAA36525.1"  
/db\_xref="GI:20037"  
/db\_xref="SWISS-PROT:P24805"  
/translation="MLAVFQSGIRPPPELSLPQAGIQKKEAKTREIAESFKTWKOD  
STFYHLENGNMFAPSHGNENPLQPSIVMDDVFCIFSGALDNTFDLRKHVGLSRQAT  
EAMIMVEAYKVLDRAPYPDPQVTKLEKGFALFLEDSKASTLFAR"  
complement(3076..4046)  
/gene="TSJT1"  
/number=2  
complement(4047..4287)  
/gene="TSJT1"  
/number=2  
complement(join(4047..4287,4473..>4678))  
/gene="TSJT1"  
/product="stem specific, weakly expressed in other organs"  
complement(4288..4472)  
/gene="TSJT1"  
/number=1  
complement(4473..>4678)  
/gene="TSJT1"  
/number=1  
complement(4778..4785)  
/note="TA29"  
BASE COUNT 2042 a 1082 c 1120 g 2010 t

0.

Db	574	ATTTAAATATTTTTTACTACAGTGCCTTGGAGTAATAATGGTGTGGAGTATGCTGTAG	511
Qy	4381	aatgtttatggaagaaatagtaaaaggttaatatgatcaaatcttcattgctatttaattgtt	444
Db	514	AATCTTTATGTGAAGAAATAGTAAAGGCTTAATATGATCAATTTTCATTTGCTATTTAATGTT	455
Qy	4441	aaaatgtgaaatttcttaattctgtgtgaaaaaacaccccaaaaaaacacttattgtggacgga	450
Db	454	AAAATGTGCAATTTCTTAATCTGTGTGAAAAACAACCAAAAAATCACTATTGTGGACCGGA	395
Qy	4501	gaaagtataaaatataatatttggaaagcgactaaaaataaaactttctcatatatacga	456
Db	394	GAAAGTATATAATATATATTTTGGACGCACTAAAAATAAACTTTTCTCATATATTACGA	335
Qy	4561	acctaaaaacagcatatgttagtttctagggaatcctaaatcactaaaaataaaaaagaa	462
Db	334	ACCTAAAAACAGCATATGTTTCTTAGGGAATCTAAATCACTAAAAATTAATAAAGAA	275
Qy	4621	gcacaagtatcaatacatatgatttaccacggtcaaacacgaaattcgtaaatattta	468
Db	274	GCAACAAGTATCAATACATATGATTTTACCGGTCAACACGAAATTCGTAAATATTAAAT	215
Qy	4681	ataataagaattaatccaataagcctccaccctataaactaaactaaaaataaacacgc	474
Db	214	ATAATAAGAAATTAATCCAATAGCTCCCAACCTTAACCTAAAAATAAACACGC	155
Qy	4741	gaatgtattatgatcatatttatattataatgaatgtgtataatcatgtataatcaatgt	480
Db	154	GAATGTATATTATATGATCAATTTATATATTAAATGTATATCAATCATGTATATCAATGT	95
Qy	4801	ataatctatgtatgtttgagaaaaagtaaacataataatagccggctatttgggtaaa	486
Db	94	ATAATCTATGTATATGTTTGAATAAAGTAAACAATTAATATAGCCGGCTATTGTGTAAA	35
Qy	4861	aatccctaataataatcg 4877	
Db	34	AATCCCTAATAATAATCG 18	
RESULT	12		
A23334			
LOCUS	A23334	5620 bp	DNA
DEFINITION	PVE108 plasmid DNA.		
ACCESSION	A23334		
VERSION	A23334.1	GI:904321	
KEYWORDS			
SOURCE	Escherichia coli.		
ORGANISM	Plasmid Escherichia coli		
REFERENCE	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae		
AUTHORS	1 (bases 1 to 5620)		
TITLE	STAMEN-SPECIFIC PROMOTERS FROM RICE		
JOURNAL	Patent: WO 9213956-A 9 20-AUG-1992;		
FEATURES	Location/Qualifiers		
source	1..5620		
	/organism="Escherichia coli"		
	/plasmid="pVE108"		
	/db_xref="taxon:562"		
misc_feature	1..395		
	/note="pUC18 derived sequence"		
misc_feature	396..802		
	/note="Agrobacterium T-DNA derived sequence; contains		
	polyA site from napaline gene"		
promoter	1139..1683		
	/note="tapetum-specific; Nicotiana tabacum derived		
	sequence"		
promoter	1684..2516		
	/note="35S3; Califlower mosaic virus isolate CabBB-J1"		
gene	derived sequence"		
	2517..3068		
	/gene="bar"		
CDS	2517..3068		

```
/gene="bar"
/codon_start=1
/transl_table=1
/product="Phosphinotricin transferase"
/protein_id="CA01674.1"
/db_xref="GI:904322"
/translotion="MDPERRPADIRRATEADMPVCTIVNHYIETSVNRTBPQPO
EWTDLVRLRERYPWLVAEVDGEVAGIAYAGPKARNAYDWTAEISYVSPRQGL
GSTLYTHLLKSLAQGPKSVAVIIGLPNDPSVRHEALGYAPRGMLRAAGFKIGNHHD
VGFWOLFSLPVPVPRPVPVTEI"
misc_feature
3069..3356
/note="Agrobacterium T-DNA derived sequence; contains
polyA site from nopaline synthase gene"
3357..5620
misc_feature
/note="pUC18 derived sequence"
BASE COUNT 1498 a 1359 c 1329 g 1430 t 4 others
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
2574
2575
2576
2577
2578
2579
2580
2581
2582
2583
2584
2585
2586
2587
2588
2589
2590
2591
2592
2593
2594
2595
2596
2597
2598
2599
2600
2601
2602
2603
2604
2605
2606
2607
2608
2609
2610
2611
2612
2613
2614
2615
2616
```





```
Db 978 CGCGGATGCTTTTCCCGGAGCGACGCTGCAAGGTTCCCTTTTGATGACCACCGCG 1037
Qy 3268 agggcttgctctgtgatttttgtaattatcatcagtagcttatgatattgtctgaaga 3327
Db 1038 AGGGCTTGTGCTTCGATGTTTGTATGTTAAATATCAGGTAGCTTATGATATGCTGAAGA 1097
Qy 3328 taatccgcaaccccgctcaaacgctgttgataaccggtacaccatggtagcttaatttttaag 3387
Db 1098 TAATCCGCAACCCCGTCAACCGTGTGATAACCGGTACCATGGTATGCTTAATTTCTTTAAG 1157
Qy 3388 taataacttgatttgatgattgattgttactgttcaacttgcaacacagggcatata 3447
Db 1158 TAAAACTTTGATTGAGTGATGATGTTGTACTGTTACACTTGCACACAGGCGATATA 1217
Qy 3448 tagagcaagacatacacacaacttgcaaaactaacttttggtagcatttcgagga 3507
Db 1218 TAGAGCAACACACATACACAACTTGCAAAACTAACCTTTGTTGGAGCATTTTCGAGGA 1277
Qy 3508 aaatggggagtagcaggctaatctgagggtaaacattaaaggtttcaattttgttg 3567
Db 1278 AAATGGGGAGTAGCAGGCTAATCTGAGGGTAACATTAAGGTTTCATGATTAATTTGTTG 1337
Qy 3568 caaacatggacttagtgagggaaagtagtaacaaattttgtctcaccctgatttcagtt 3627
Db 1338 CAAACATGGACTTAGTGTGAGGAAAAAGTACCAAAATTTTGTCTCACCCTGATTTTCAGTT 1397
Qy 3628 atggaattacattatgaagctgtgtagagaagatgtttattctagtcagccacccac 3687
Db 1398 ATGGAATTACATTTATGAAGCTGCTAGAGAAAGATGTTTATCTAGTCCAGCCACCCAC 1457
Qy 3688 ctatgcagctcgtctttagctgattgattcaaaactgatttaattacattgttaaatgt 3747
Db 1458 CTATGCAAGCTGCTTTTACGTGATGATTCAAAACACTGATTAATTAATGCTAAATGT 1517
Qy 3748 gcatactcgagcctatgtcgtcttaattcagtagtagatgtttattctagtcagacataaaaa 3807
Db 1518 GCATACTTTCGAGCTATGTCGCTTTAAATCGAGTAGATGATATATTAGTACATAAAAA 1577
Qy 3808 atcatgttgaaatcaacttttcaataaagtgacaaagtcaattgtcccttctgttgacact 3867
Db 1578 ATCATGTTTGAATCAATCTTTCATPAAAGTGACAAAGTCAAAATGTGCTTCTGTTGGCAGT 1637
Qy 3868 atattcaactgttaattgcaaatccagttataacttagctagat 3913
Db 1638 ATATTCAATCTGTTAATGCAAAATTATCCAGTTATACCTTAGCTAGAT 1683

RESULT 15
LOCUS I47737 5620 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 9 from patent US 5639948.
ACCESSION I47737
VERSION I47737.1 GI:2471702
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1. (bases 1 to 5620)
AUTHORS Michiels,F., Morioka,S., Scheirlinck,T. and Komari,T.
TITLE Stamen-specific promoters from rice
JOURNAL Patent: US 5639948-A 9 17-JUN-1997;
FEATURES Location/Qualifiers
1. 5620
/organism="unknown"
BASE COUNT 1498 a 1359 c 1329 g 1430 t 4 others
ORIGIN
```

Query Match 21.7%; Score 1270.4; DB 10; Length 5620;  
Best Local Similarity 98.9%; Pred. No. 6.1e-179;  
Matches 1291; Conservative 0; Mismatches 11; Indels 4; Gaps 1;

```
Qy 2612 ttctgaccgggtaccgggaattcaggctcgggtaccggggatcttcccgcattagtaac 2671
Db 378 TGTAAACAGACGGCCAGTGAATTCGAGCTCGGTACCGGGATCTTCCCGATCTAGTAAC 437
Qy 2672 atagatgaccccgccgcgcgaatttatctcctagtttgcgcgctatatcttggtttttatc 2731
Db 438 ATAGATGACACCGCGCGGATAAATTTATCTAGTTTTCGCGCTATATTTTGTGTTTCTATC 497
Qy 2732 gcgtattaaatgtataattgcgggactcctaatacataaaaaccccatctcataaataacgctc 2791
Db 498 GCCTATTAAATGTATAATTCGGGACTCTAATCATATAAAACCCATCTCATATAAATAACGTC 557
Qy 2792 atcattataatgttaattattacatgcttaacgttaattcaacagaataattatataaatac 2851
Db 558 ATGCATTACATGTTAATTAATTACATGCTTAAACGTAATTCACAGAAATATATATGATAATC 617
Qy 2852 atcgaagacggcgaacaggattcaacttaagaactttattgcccgaatgttgaacga 2911
Db 618 ATCGCAAGACCGGCAACAGATTCATCTTAAGAAACTTTTATGCCAAATGTTTGAACA 677
Qy 2912 tctgcttggaatcctctagag----ccggaagtgaaattgaccgatcagagtttgaaga 2967
Db 678 TCTGCTTCGGATCCTCTAGAGNNNCCGAAAGTGAATTTGACCAACTTTTATGCCAAATGTTTGAACA 737
Qy 2968 aaaaattattacacactttatgataagctgaaaaaacggcctccgcaggaagccgtttt 3027
Db 738 AAAATTTATTACACACTTTATGTAAAGCTGAAAAAACGGCTCCGCAAGGAAGCGGTTTT 797
Qy 3028 ttctgttatctgattttttaaaggctctgataaattggtccgttggtttggtaaatcagccag 3087
Db 798 TTTTCGTTATCTCATTTTGTAAAGTCTGATTAATGGTCCGTGTTTGTAAATATCCGCT 857
Qy 3088 tcgcttgagtaagaatccggctcgtgaattctcgaagcctgatgtatagttaatatccgct 3147
Db 858 TCGCTTGTAGTAAGAAATCCGGTCTGAATTTCTGAAGCCTGATGATAGTAAATATCCGCT 917
Qy 3148 tcaacccatgttcgctccgcttttgcgggagtttgccttccctgttttgagaagaatgct 3207
Db 918 TCACGCCATGTTTCGTCGCTTTTGGCCGGAGTTTGGCCTTCCCTGTTTGTGAGAAGATGCT 977
Qy 3208 ccgcgatgcttttccccggagcgactcgtcaagggttcccttttggaccaccagccg 3267
Db 978 CGCGCGATGCTTTTCCCGGAGCGACGCTGCAAGGTTCCCTTTTGTAGCCACCCAGCGC 1037
Qy 3268 agggcttgctctgattttgtaattatcattatcaggtagcttatgatattgtctgaaga 3327
Db 1038 AGGGCTTGTGCTTCGATTTTGTAAATGTAATATCAGGTAGCTTATGATATGCTGAAGA 1097
Qy 3328 taatccgcaaccccgctcaaacgctgttgataaccggtacacatggtagcttaatttttaag 3387
Db 1098 TAATCCGCAACCCCGTCAACCGTGTGATAAACCAGTACCATGGTAGCTAATTTCTTTAAG 1157
Qy 3388 taataacttgatttgatgattgattgttactgttcaacttgcaacacagggcatata 3447
Db 1158 TAAAACTTTGATTGAGTGATGATGTTGTACTGTTACACTTGCACACAGGCGATATA 1217
Qy 3448 tagagcaagacatacacacaacttgcaaaactaacttttggtagcatttcgagga 3507
Db 1218 TAGAGCAACACACATACACAACTTGCAAAACTAACCTTTGTTGGAGCATTTTCGAGGA 1277
Qy 3508 aaatggggagtagcaggctaatctgagggtaaacattaaaggtttcaattttgttg 3567
Db 1278 AAATGGGGAGTAGCAGGCTAATCTGAGGGAACATTAAGGTTTTCATGATTAATTTCTTGT 1337
Qy 3568 caaacatggacttagtgagggaaagtagtaacaaattttgtctcaccctgatttcagtt 3627
Db 1338 CAAACATGGACTTAGTGTGAGGAAAAAGTACCAAAATTTTGTCTCACCCTGATTTTCAGTT 1397
Qy 3628 atggaattacattatgaagctgtgtagagaagatgtttattctagtcagccacccac 3687
Db 1398 ATGGAATTACATTTATGAAGCTGCTAGAGAAAGATGTTTATCTAGTCCAGCCACCCAC 1457
Qy 3688 ctatgcagctcgtctttagctgattgattcaaaactgatttaattacattgttaaatgt 3747
```

Db 1458 CTTATGCAAGTCGCTTTTATGCTTGATTCAAAACTGATTAAATTACATTTGCTAAATGT 1517

Qy 3748 gcatacttcgagcctatgtcgctttaattcagtaggatgtatatattagtcacataaaaa 3807

Db 1518 GCATACTTCGAGCCTATGTCGCTTTAATTCGAGTAGGATGTATATATTAGTACATAAAAA 1577

Qy 3808 atcatgtttgaatcatcttttcataaaagtgacaagtcgaattgtcccttccttggttggcaact 3867

Db 1578 ATCATGTTTGAATCATCTTTTCATAAAAGTGACAAGTCAATTGTCCCTTCTTGTGTTGGCACT 1637

Qy 3868 atattcaatctgttaataatgcaaaattatccagttatacttagctagat 3913

Db 1638 ATATTCARTCTGTATATGCAAAATTATCCAGTTATACCTAGCTAGAT 1683

Search completed: November 15, 2001, 20:39:18

Job time: 15545 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2001, 16:23:38 ; Search time 328.78 Seconds  
(without alignments)  
11200.934 Million cell updates/sec

Title: US-09-698-903B-1

Perfect score: 5865

Sequence: 1 aattacaacggtatatcc.....caattgaatatcctgcg 5865

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_0601.\*  
1: /SID88/gcgdata/geneseq/geneseqn/NA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseqn/NA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseqn/NA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseqn/NA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseqn/NA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseqn/NA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseqn/NA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseqn/NA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseqn/NA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseqn/NA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseqn/NA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseqn/NA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseqn/NA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseqn/NA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseqn/NA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseqn/NA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseqn/NA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseqn/NA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseqn/NA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseqn/NA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseqn/NA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseqn/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	5849.8	99.7	5864	17 AAT39339	Plasmid pTCO113 T-
2	4903.2	83.6	4946	18 AAT59531	T-DNA of plasmid p
3	1500.2	25.6	3265	10 AA02188	TA29 gene insert i
4	1268.8	21.6	5620	13 AAQ27489	pVE108. Synthetic
5	1267.2	21.6	5620	15 AAQ53875	Plasmid pVE108 use
6	885.8	15.1	5349	19 AAV23239	T-DNA of pTTS24.
7	832.8	14.2	6539	17 AAT391097	E. coli plasmid pT
8	832.8	14.2	6548	17 AAT39336	Plasmid pTTS174 use
9	832.8	14.2	6548	18 AAT61394	Plasmid pTTS172. C
10	832.8	14.2	6548	21 AAZ91096	E. coli plasmid pT
11	802.2	13.7	7599	22 AAF25320	Nucleotide sequenc

12	741.8	12.6	3658	17 AAT39338	Plasmid pVE136 (Ec
13	727.8	12.4	4808	15 AAQ53889	Restriction fragme
c 14	712	12.1	4896	17 AAT08976	pTS256 contg. p35S
c 15	708.8	12.1	4896	20 AAX15631	Nucleotide sequenc
16	706.8	12.1	6555	15 AAQ53874	Plasmid pVE144 use
17	705.4	12.0	6376	13 AAQ27951	pVE149. Escherich
18	595.2	10.1	1303	17 AAQ39337	Plasmid pTS88 (eco
c 19	564	9.6	790	17 AAT13117	Inserted DNA fragm
c 20	564	9.6	790	21 AAZ39903	DNA fragment from
c 21	561.6	9.6	2345	19 AAV54016	Nucleotide sequenc
c 22	561.6	9.6	2345	20 AAX08450	T-DNA sequence of
c 23	561.6	9.6	2345	22 AAC58664	Agrobacterium tume
c 24	561.6	9.6	3544	17 AAT17246	pTS200 contg. p35S
c 25	561.6	9.6	3544	20 AAX15632	Nucleotide sequenc
c 26	561.6	9.6	3658	17 AAT39338	Plasmid pVE136 (Ec
c 27	561.6	9.6	4808	15 AAQ53889	Restriction fragme
c 28	561.6	9.6	4883	13 AAQ25709	pDE110. Synthetic
c 29	561.6	9.6	4883	14 AAQ51192	pDE110; plasmid DN
30	561.6	9.6	4896	17 AAT08976	pTS256 contg. p35S
31	561.6	9.6	4896	20 AAX15631	Nucleotide sequenc
32	561.6	9.6	5611	19 AAV23242	Plasmid pLH48. Sy
c 33	561.6	9.6	5620	13 AAQ27489	pVE108. Synthetic
c 34	560	9.5	6376	13 AAQ27951	pVE149. Escherich
c 35	558.4	9.5	5620	15 AAQ53875	Plasmid pVE108 use
c 36	551.8	9.4	9335	20 AAV63734	Plasmid pGPG/367 c
c 37	549.6	9.4	610	21 AAZ61377	Nucleotide sequenc
c 38	549.6	9.4	838	8 AAT71305	Sequence encoding
c 39	549	9.4	549	13 AAQ23354	Male organ-specifi
c 40	549	9.4	5611	19 AAV23242	Plasmid pLH48. Sy
c 41	548.8	9.4	552	18 AAT61880	Streptomyces hygro
42	491.4	8.4	6548	17 AAT39336	Plasmid pTTS174 use
43	491.4	8.4	6548	18 AAT61394	Plasmid pTTS172. C
44	491.4	8.4	6548	21 AAZ91096	E. coli plasmid pT
45	487.2	8.3	6539	21 AAZ91097	E. coli plasmid pT

#### ALIGNMENTS

RESULT 1  
AAT39339  
ID AAT39339 standard; DNA; 5864 BP.  
XX  
AC AAT39339;  
XX  
DT 22-JAN-1997 (first entry)  
XX  
DE Plasmid pTCO113 T-DNA used to obtain male sterile oilseed rape.  
XX  
KW Plasmid pTCO113; male sterile; barnase; ribonuclease; barstar;  
transgenic plant; oilseed rape; canole; Brassica napus; ds.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT misc\_feature complement (1..25)  
FT /tag= a  
FT /label= RB  
FT /note= "right border of Agrobacterium T-DNA"  
FT polyA\_signal complement (98..330)  
FT /tag= b  
FT /label= 3'g7  
FT /note= "region containing polyA signal of gene 7  
of Agrobacterium T-DNA"  
FT CDS complement (331..882)  
FT /tag= c  
FT /label= bar  
FT /note= "region coding for phosphinothricin  
acetyltransferase"  
FT promoter complement (883..2608)  
FT /tag= d  
FT /label= Pssu  
FT /note= "promoter of Arabidopsis Rubisco small











. QY 301 agctcatcgggggatcctagaacgcgtgatctcagatctcggtagacggcgacgccggac 360  
Db 301 agctcatcgggggatcctctag-acgcgtgagatcagatctcggtagacggcgacgccggac 359  
QY 361 gggcgctaccggcagcgtgaagtcagctgcagaaacccacgctcatgccaagtcccgct 420  
Db 360 gggcgctaccggcagcgtgaagtcagctgcagaaacccacgctcatgccaagtcccgct 419  
QY 421 gcttgaagccggccgcccgcagcatgcccggggggcataatccagcgcctcgtgcacgc 480  
Db 420 gcttgaagccggccgcccgcagcatgcccggggggcataatccagcgcctcgtgcacgc 479  
QY 481 gcaagctcgggtcgcttgggagcccgatgacagcgaccacgctcttgaagccctgtgcct 540  
Db 480 gcaagctcgggtcgcttgggagcccgatgacagcgaccacgctcttgaagccctgtgcct 539  
QY 541 ccagggaacttcagcaggctgggtgtagagcgtgagcccagtcctcgcgtggtggcggg 600  
Db 540 ccagggaacttcagcaggctgggtgtagagcgtgagcccagtcctcgcgtggtggcggg 599  
QY 601 gggagacgtacacggtcgaactcggccgtccagtcgtagcgttggtgcctccaggggc 660  
Db 600 gggagacgtacacggtcgaactcggccgtccagtcgtagcgttggtgcctccaggggc 659  
QY 661 ccgcgtaggcgtgcgcgcacccctgcgcgtccacgtcggcagcagccagggatagcgt 720  
Db 660 ccgcgtaggcgtgcgcgcacccctgcgcgtccacgtcggcagcagccagggatagcgt 719  
QY 721 cccgcagcggacgaggctcgctccgctccactcctcgggttcctgcggtcgtgacggaggt 780  
Db 720 cccgcagcggacgaggctcgctccgctccactcctcgggttcctgcggtcgtgacggaggt 779  
QY 781 tgaccgtctctcgtatgtagtggttgacgatgggtgacagcccgccgcatgccct 840  
Db 780 tgaccgtctctcgtatgtagtggttgacgatgggtgacagcccgccgcatgccct 839  
QY 841 cggttggcagccgggatgcggccggggcgctcgttggttccattgtctctcttactctt 900  
Db 840 cggttggcagccgggatgcggccggggcgctcgttggttccattgtctctcttactctt 899  
QY 901 tgtgtgactcagggttgggtctagtcgtttggctatctatatataatgatacaacaatga 960  
Db 900 tgtgtgactcagggttgggtctagtcgtttggctatctatatataatgatacaacaatga 959  
QY 961 gaacaaagctttggagtgatcggagggctcaggatcacatgagattcaagtggactagatc 1020  
Db 960 gaacaaagctttggagtgatcggagggctcaggatcacatgagattcaagtggactagatc 1019  
QY 1021 tacacgcttggatttgaagtggatagtgtgagggttaattttacttggtaacggccac 1080  
Db 1020 tacacgcttggatttgaagtggatagtgtgagggttaattttacttggtaacggccac 1079  
QY 1081 aaaggccaaaggagaggtgttgagacccttatcggtctgaaccgctggaataatgccag 1140  
Db 1080 aaaggccaaaggagaggtgttgagacccttatcggtctgaaccgctggaataatgccag 1139  
QY 1141 tggagaataatccatgaattctatcgttatctatgagtgaaattgtgtgagtgagggt 1200  
Db 1140 tggagaataatccatgaattctatcgttatctatgagtgaaattgtgtgagtgagggt 1199  
QY 1201 ggtgcttgctcatcttacttgccgtggacttggcccttccattatggtgggaatttatat 1260  
Db 1200 ggtgcttgctcatcttacttgccgtggacttggcccttccattatggtgggaatttatat 1259  
QY 1261 tttaacttactatagagctttcaacaccttttttttaacctggatttgaattatataatg 1320  
Db 1260 tttaacttactatagagctttcaacaccttttttttaacctggatttgaattatataatg 1319  
QY 1321 gtagtattcatgaataaaaaatgggaatttttgaatttgaactgctaaatgcataagatt 1380  
Db 1320 gtagtattcatgaataaaaaatgggaatttttgaatttgaactgctaaatgcataagatt 1379  
QY 1381 aggtgaaactgtgggaataataatttttttcaatttaaaagcaaaatttgccttttactaga 1440

. Db 1380 aggtgaaactgtgggaataataattttttcatttaaaagcaaaatttgccttttactaga 1439  
QY 1441 attataaataatagaataataataataataataataataataataataataataataata 1500  
Db 1440 attataaataatagaataataataataataataataataataataataataataataata 1499  
QY 1501 aacgagaactatgttttaattgtttaaagattagtcgacatcagtcacatcgtttacaatatg 1560  
Db 1500 aacgagaactatgttttaattgtttaaagattagtcgacatcagtcacatcgtttacaatatg 1559  
QY 1561 ttacaacaagtcataagcccacaagaattagcagctctataataaactaaagagtcacacga 1620  
Db 1560 ttacaacaagtcataagcccacaagaattagcagctctataataaactaaagagtcacacga 1619  
QY 1621 aaatatcacaaatcataagcccacaagaattagcagctctataataaactaaagagtcacacga 1680  
Db 1620 aaatatcacaaatcataagcccacaagaattagcagctctataataaactaaagagtcacacga 1679  
QY 1681 agctaaacaaagtcocaaaaaaacttctcaagtcctccatctctctttatgaacattgaaa 1740  
Db 1680 agctaaacaaagtcocaaaaaaacttctcaagtcctccatctctctttatgaacattgaaa 1739  
QY 1741 actatacacaaaaaagtcagataaaatctcttctggcgtgcttcccaaacctcttaca 1800  
Db 1740 actatacacaaaaaagtcagataaaatctcttctggcgtgcttcccaaacctcttaca 1799  
QY 1801 tcaactccctatcggattggaattgttttacttctgacatttccgttgcgaatgatattgata 1860  
Db 1800 tcaactccctatcggattggaattgttttacttctgacatttccgttgcgaatgatattgata 1859  
QY 1861 gtatgtttgtgaacactaaatagggtttaacaatcgaagtcagtggaatcattggattgttgc 1920  
Db 1860 gtatgtttgtgaacactaaatagggtttaacaatcgaagtcagtggaatcattggattgttgc 1919  
QY 1921 agattttccgagagcgttcttagtagaaaagcccatccacgaaatttactagtaaaataaa 1980  
Db 1920 agattttccgagagcgttcttagtagaaaagcccatccacgaaatttactagtaaaataaa 1979  
QY 1981 tcaccaattaggtttcttattatgtgccaaattccaataataattatagaggaattttcaca 2040  
Db 1980 tcaccaattaggtttcttattatgtgccaaattccaataataattatagaggaattttcaca 2039  
QY 2041 tgaacacgtatgaattgtattagtaaaatggtcaggttaagacattaaaaaaatccctacgct 2100  
Db 2040 tgaacacgtatgaattgtattagtaaaatggtcaggttaagacattaaaaaaatccctacgct 2099  
QY 2101 agatatccaactttaaaaaattcgaatcagtcgtggaattgtacaaaaatttgggatctacta 2160  
Db 2100 agatatccaactttaaaaaattcgaatcagtcgtggaattgtacaaaaatttgggatctacta 2159  
QY 2161 tatataataatgctttacacacacttggatttttttggagcgtggaatttttactata 2220  
Db 2160 tatataataatgctttacacacacttggatttttttggagcgtggaatttttactata 2219  
QY 2221 catatttgtttggccatgcacacaactcattgttttagtgttaactttgtatttggatcaca 2280  
Db 2220 catatttgtttggccatgcacacaactcattgttttagtgttaactttgtatttggatcaca 2279  
QY 2281 tatatgttgcgtgtatatttataaagaatttcttggaccataacacacacacata 2340  
Db 2280 tatatgttgcgtgtatatttataaagaatttcttggaccataacacacacacata 2339  
QY 2341 tatataataataattatataatcatcgaactttttaaattgaaaaaaataatataatat 2400  
Db 2340 tatataataataattatataatcatcgaactttttaaattgaaaaaaataatataatat 2399  
QY 2401 atagtgcatttttttaacaacacataatgttggcgtgattcgtcgaataatcactgctaga 2460  
Db 2400 atagtgcatttttttaacaacacataatgttggcgtgattcgtcgaataatcactgctaga 2459  
QY 2461 gtaatgaaaaataataatctattgtggaattatctccagatgttaagatttttcttaagata 2520

Db	2460	gtaatgaaaaataataatctattgtctgaaattatctcagatggttaagattttctttaagta	2519
QY	2521	aattcttcaaaattttagctaaaagctctgttaataactaaagaataatacacacatctcga	2580
Db	2520	aattcttcaaaattttagctaaaagctctgttaataactaaagaataatacacacatctcga	2579
QY	2581	ccacgaaaaaaacacataataaatttgaatttcgacgcggtaccccggaattcgagct	2640
Db	2580	ccacgaaaaaaacacataataaatttgaatttcgacgcggtaccccggaattcgagct	2639
QY	2641	cggtaccggggatctcccgatctagtaacatagatgacacgcgcgcgataatttacc	2700
Db	2640	cggtaccggggatctcccgatctagtaacatagatgacacgcgcgcgataatttacc	2699
QY	2701	ctagtttgcgcgtatatattttgttttctatcgcggtattaaattgataattgcgggactct	2760
Db	2700	ctagtttgcgcgtatatattttgttttctatcgcggtattaaattgataattgcgggactct	2759
QY	2761	aatcataaaaaaccatctcataaataacgctcatgcattacatgtttaatttaccatgctt	2820
Db	2760	aatcataaaaaaccatctcataaataacgctcatgcattacatgtttaatttaccatgctt	2819
QY	2821	aacgtaattcaacagaaattatgatataatcatcgcgaagccgcgaacagattcaatct	2880
Db	2820	aacgtaattcaacagaaattatgatataatcatcgcgaagccgcgaacagattcaatct	2879
QY	2881	taagaaactttattgcocaaatgtttgaacgatctgttcggatccctctagagcggaaaag	2940
Db	2880	taagaaactttattgcocaaatgtttgaacgatctgttcggatccctctagagcggaaaag	2939
QY	2941	tgaattgaccgatcagagtgttgaagaaaaatttattacacattttatgtaaaagctgaaa	3000
Db	2940	tgaattgaccgatcagagtgttgaagaaaaatttattacacattttatgtaaaagctgaaa	2999
QY	3001	aaacggctcccgaggaagccgttttttcgttatctgatttttgtaaaaggtctgataa	3060
Db	3000	aaacggctcccgaggaagccgttttttcgttatctgatttttgtaaaaggtctgataa	3059
QY	3061	tggctcgtgtttttgtaaaacagccagctgcttgagtaaaagaatccggtctgtaattctg	3120
Db	3060	tggctcgtgtttttgtaaaacagccagctgcttgagtaaaagaatccggtctgtaattctg	3119
QY	3121	aagcctgatgcatagtttaataatccgcttcacgcacatgttctgctcgtctttgcccggaggt	3180
Db	3120	aagcctgatgcatagtttaataatccgcttcacgcacatgttctgctcgtctttgcccggaggt	3179
QY	3181	ttgccttccctgtttgagaagatgtctccgcgatgcttttccccggagcagctctgca	3240
Db	3180	ttgccttccctgtttgagaagatgtctccgcgatgcttttccccggagcagctctgca	3239
QY	3241	aggttcccttttggatgcaccccgagggcgttggcttctgattttgtaagttaatta	3300
Db	3240	aggttcccttttggatgcaccccgagggcgttggcttctgattttgtaagttaatta	3299
QY	3301	tcagggtagcttcatgatatgtctgaagataatccgcgaaccocgctcaaacgcttgataacc	3360
Db	3300	tcagggtagcttcatgatatgtctgaagataatccgcgaaccocgctcaaacgcttgataacc	3359
QY	3361	ggtaccatggtagctaatttctttaagtaaaactttgatttttgagtgatgattgttact	3420
Db	3360	ggtaccatggtagctaatttctttaagtaaaactttgatttttgagtgatgattgttact	3419
QY	3421	gttacacttgcacacaaggccataatagagcacagacatacacacaacacttgcaaaa	3480
Db	3420	gttacacttgcacacaaggccataatagagcacagacatacacacaacacttgcaaaa	3479
QY	3481	ctaacctttgttgagcatcttcaggaaaaatggggagtagcaggcttaacttgaggggtaac	3540
Db	3480	ctaacctttgttgagcatcttcaggaaaaatggggagtagcaggcttaacttgaggggtaac	3539
QY	3541	attaaaggtttcatgattaatttgtgcaaacatggacttagtgcgaggaagaaagtacca	3600
Db	3540	attaaaggtttcatgattaatttgtgcaaacatggacttagtgcgaggaagaaagtacca	3599
QY	3601	aaattttgtctcccttgatttcagtttatggaaattacattatgaagctgtgctagagaa	3660
Db	3600	aaattttgtctcccttgatttcagtttatggaaattacattatgaagctgtgctagagaa	3659
QY	3661	gatgtttattctctagtcagccacccacccttatgcgaagctgcgttttagcttgattcaaaa	3720
Db	3660	gatgtttattctctagtcagccacccacccttatgcgaagctgcgttttagcttgattcaaaa	3719
QY	3721	acgattttaattcacattgctaaaatgctacacttcgagcctatgctgctttaaattcgag	3780
Db	3720	acgattttaattcacattgctaaaatgctacacttcgagcctatgctgctttaaattcgag	3779
QY	3781	taggatgtatatatttagtacataaaaaatcatgtttgaaatcatcttttcaataaagtgacaa	3840
Db	3780	taggatgtatatatttagtacataaaaaatcatgtttgaaatcatcttttcaataaagtgacaa	3839
QY	3841	gtcaattgtcccttcttctgtttggcactatattcaactgtttaatgtcaaaattatccaagta	3900
Db	3840	gtcaattgtcccttcttctgtttggcactatattcaactgtttaatgtcaaaattatccaagta	3899
QY	3901	tacttagctagatatccaatttttgaataaaaaatagctctcttgattagtaaaacccgagatgtg	3960
Db	3900	tacttagctagatatccaatttttgaataaaaaatagctctctcttgattagtaaaacccgagatgtg	3959
QY	3961	acaaagtccacatatccaatcaaaacttctggtgctgctggctaaagtcttgatcgacatgggg	4020
Db	3960	acaaagtccacatatccaatcaaaacttctggtgctgctggctaaagtcttgatcgacatgggg	4019
QY	4021	ttaaaatttaaaattgggacacataaaatagctcttggtaagtcttgatcgacatgggg	4080
Db	4020	ttaaaatttaaaattgggacacataaaatagctcttggtaagtcttgatcgacatgggg	4079
QY	4081	acagattggtcacatggaacacaaaaagctcctctgataagaagtcgcaaaagtatcacaaattt	4140
Db	4080	acagattggtcacatggaacacaaaaagctcctctgataagaagtcgcaaaagtatcacaaattt	4139
QY	4141	tcctatcgagagatagattgaaagaagtgcagggaagcgggttaactggaacataacacaat	4200
Db	4140	tcctatcgagagatagattgaaagaagtgcagggaagcgggttaactggaacataacacaat	4199
QY	4201	gtctaaattaatgtcattctgcttaacccaaaaagtgatatctactctctccgggtcccaataag	4260
Db	4200	gtctaaattaatgtcattctgcttaacccaaaaagtgatatctactctctccgggtcccaataag	4259
QY	4261	ttatttttggcccttttttatgtgtccaaaaataagtgtgttttttagatttcaaaaaatg	4320
Db	4260	ttatttttggcccttttttatgtgtccaaaaataagtgtgttttttagatttcaaaaaatg	4319
QY	4321	atttaattattttttactacagtgcccttggagtaaatgggtgttggagtatgttgttaga	4380
Db	4320	atttaattattttttactacagtgcccttggagtaaatgggtgttggagtatgttgttaga	4379
QY	4381	aatgtttatgtgaagaaatagtaaaaggttaatatgatcaaatttcoatttgcattttaagttt	4440
Db	4380	aatgtttatgtgaagaaatagtaaaaggttaatatgatcaaatttcoatttgcattttaagttt	4439
QY	4441	aaaaatgtaatttctttaaactgtgtgaaaaacacacaaaaaacacttattgtggaccgga	4500
Db	4440	aaaaatgtaatttctttaaactgtgtgaaaaacacacaaaaaacacttattgtggaccgga	4499
QY	4501	gaaagtataaaaaatataattttggaagcgcactaaaaataaaccttttctcatattatacga	4560
Db	4500	gaaagtataaaaaatataattttggaagcgcactaaaaataaaccttttctcatattatacga	4559
QY	4561	acctaaaaacagacataggttagtttcttagggaaatcttaaaatcaactaaaaatataaaagaa	4620
Db	4560	acctaaaaacagacataggttagtttcttagggaaatcttaaaatcaactaaaaatataaaagaa	4619
QY	4621	gcaacaagttaacatcatatgatttacacgcgtcaaacacgaaatctcgtaaaatttaatt	4680
Db	4620	gcaacaagttaacatcatatgatttacacgcgtcaaacacgaaatctcgtaaaatttaatt	4679

Qy	4681	ataataagaattaatccaaatagctcccacccctataacttaaacataaaccaggc	4740
Db	4680	ataataagaattaatccaaatagctcccacccctataacttaaacataaaccaggc	4739
Qy	4741	gaatgtatatcatgcataatttatatatattaaaatggtataatcaatgatgt	4800
Db	4740	gaatgtatatcatgcataatttatatatattaaaatggtataatcaatgatgt	4799
Qy	4801	ataatcctatgtatatgggttagaaaaagttaaacaatttaatatagccggctattt	4860
Db	4800	ataatcctatgtatatgggttagaaaaagttaaacaatttaatatagccggctattt	4859
Qy	4861	aatccctaataatcgcgacggatccccgggaattccggggaagcttagatccatgcag	4920
Db	4860	aatccctaataatcgcgacggatccccgggaattccggggaagcttagatccatgcag	4919

RESULT 3  
AAN92188/c

ID AAN92188 standard; DNA; 3265 BP.

AC AAN92188;

DT 11-APR-1990 (first entry)

DE TA29 gene insert in pTA29S3.

XX TA29 promoter; pTA29S3;

KW Nicotiana tabacum "Samsun".

OS

FH Key Location/Qualifiers

FT TATA-signal 1445..1451

FT \*tag= a

FT CDS 1526..4991

FT \*tag= b

FT CDS 1940..2296

FT \*tag= c

XX

PX WO8910396-A.

PN

PD 02-NOV-1989.

XX

PF 27-APR-1989; 89WO-EP00495.

XX

PR 28-APR-1988; 88GB-0010120.

XX

PA (PLAN-) PLANT GENETIC SYSTEMS NV.

XX

PI Mariani C, Leemans J, De Greef W, De Beuckeleer M;

XX

DR WPI; 1989-339961/46.

DR P-FSDB; AAP93313.

XX

PT Transformed plant cell with disrupted metabolism - from a product encoded by foreign male-sterility DNA inserted in the nuclear genome.

XX

PS Example 2; fig. 3A; 89pp; English.

XX

CC Plasmid pTA29S3 was constructed subcloning a ClaI fragment contg. the TA29 gene from lambda TA29 into plK31. The TA29 promoter is used in a vector for transforming plant cells. Tag b is the coding sequence of the TA29 gene; tag c is the cDNA insert from pBR329 (AAN92187).

XX

SQ Sequence 3265 BP; 1016 A; 581 C; 623 G; 1045 T; 0 other;

Query Match	25.6%;	Score 1500.2;	DB 10;	Length 3265;
Best Local Similarity	99.7%;	Pred. No. 3e-242;		
Matches 1513;	Conservative	0;	Mismatches 3;	Indels 1;
Gaps	1;			

Qy	3361	ggatccatggtagctaaattttttaagtaaaaccttgattgagtgatgatgtgtact	3420
Db	3361	ggatccatggtagctaaattttttaagtaaaaccttgattgagtgatgatgtgtact	3420

Db	1533	GCTACCAATTTAGCTAAATTTCTTTAAGTAAAAACCTTTGATTTGAGTGATGATGTTGTACT	1474
QY	3421	gttacacttgcaccaaaggccatataagagcacaagacatacacacaacttcgaaaa	3480
Db	1473	GTTACACTTGCACCACAGGGCATATATAGAGCACAGACATACACAACTTTCGAAAA	1414
QY	3481	ctaaactttgttgaggaacttcgaggaaaaatggggagtagcaggctaaactctgagggtaac	3540
Db	1413	CTAACTTTTGTGTGGAGCATTTTCGAGGAAATTCGGAGTAGTACGAGGCTAATCTGAGGCTAAC	1354
QY	3541	attaagggttccatgatttaatttgttgcaaacatggaacttagtgcagggaaaaaagtacc	3600
Db	1353	ATTAAAGTTTCATGTATTATTTGTTGCAAACTATGACATAGTGTAGTGAGGAAAAAGTAGTACCA	1294
QY	3601	aaatttctccaccctgatttcagtttattgtgaaattacattatgaactgaagctgctcagagaa	3660
Db	1293	AAATTTTGTCTCACCCTGATTTTCAGTTATGGAATTTACATTAATGAAGCTGTGCTAGAGAA	1234
QY	3661	gatgttattctagtcacagccaccaccttctgcaagctgcgtcttttagcttgatttcacaaa	3720
Db	1233	GATGTTTATTTCTAGTCCAGCACCCACCTTATCGAAGTCTGCTTTAGCTTTGATTCAAAA	1174
QY	3721	actgatttcaattacattgcctaaatgctcacttcgagcctatgctgcctttaatttcog	3780
Db	1173	ACTGATTTAATTTACATTTGCTAAATGTGCATACTTCGAGCCTATGTGCTTTTAATTCGAG	1114
QY	3781	taggatgtatatattagcacataaaaaatcatgtttgaaatcatctcttcataaaagtgcac	3840
Db	1113	TAGGATGTATATATTAGTACATAAAAAATCATGTTTGAATCATCTTTCATAAAGTGACAA	1054
QY	3841	gtcaattgcccctcttctgttggcacatattccaatctgtttaaagcaaatctccagtt	3900
Db	1053	GTCAAATGTCCCTTCCTTGTGTGGCACTATATTCAACTCTGTTAAATGCAAAATATCCAGTTA	994
QY	3901	tacttagctagatccaatttttgaataaaaaatgctcttcattagtaaacccgagtagtg	3960
Db	993	TACTTAGTAGATATCCAAATTTTGAATAAAAAATACCTCTTGATTAGTAACCCGGATAGTG	934
QY	3961	acaaagtccacatccatccaactctgtgctgcgtggtgaagtctcgtatgcacatgggg	4020
Db	933	ACAAAGTCACATATCCATCAAACTTCGTGTGCTCGTGGCTAAGTCTTGATCGACATGGGG	874
QY	4021	ttaaaatttaaaattggagcacataaataagctatttgcgaatctcccacatcgaaaaatg	4080
Db	873	TTAAATTTAAATTTGGGACACATAAATAGCCCTATTTGTGCAAAATCTCCCATCGAAAAATG	814
QY	4081	acagattgttcacatggaacacaaaaagtcctctgatagaagtcgcaagatcacaaatt	4140
Db	813	ACAGATTGTTACATGTGAAACAAAAAAGTCCCTGATAGAAGTCGCAAGTATCACAAATTT	754
QY	4141	tcctacgagatagattgaagaagtagcagggagcgtttaactgtgaaacataacacaat	4200
Db	753	TCTATCGAGAGATAGATTGAAAGAAGTCAGGGAAGCGGTTAATCGGAACATAAACACAAT	694
QY	4201	gtctaaatttaattgcattcgtcaacccaaaagtgtattactctctccggtccacaataag	4260
Db	693	GTCTAAATTAATTTGCATTCGCTTAACCAAAAAGTGTATTACTCTCCCGGTCCACAATAAG	634
QY	4261	ttatttttggccctttttttatggtccaaaataagtgagtttttttagatttccaaaaatg	4320
Db	633	TTATTTTGTGGCCCTTTTATTTGTTCCAAAAATAGTGAGTGTGTTTTTAGATTTTCAAAATG	574
QY	4321	atttaatttttttttactacagtcaccttggagtaaatgggttggagtagtggttaga	4380
Db	573	ATTTAATTTATTTTACTACAGTGCCTTCGGAGTAATGCTGTGGAGTATGCTGTAGA	514
QY	4381	aatgtttatgtgaagaaatagtaaaaggttaatatgatccaatttcattgctatttaattgt	4440
Db	513	AATGTTTATGTGAAGAAATAGTAAGGTTAATATGATCAATTTTCATTTGCTATTTTAATGTT	454
QY	4441	aaaatgtgaattcttataatctgtgaaacacacaaaaataactatttggaccgga	4500
Db	453	AAAATGTGAATTTCTTAATCTGTGAAACACACAAAAATCACATTTATTTGTGGACCGGA	394









```

FT /label= bar
FT /note= "region coding for phosphinothricin acetyl
FT transferase"
FT complement (884..2258)
FT *tag= d
FT /label= P35S
FT /note= "35S promoter of Cauliflower Mosaic Virus"
FT 2281..3969
FT *tag= e
FT /label= PE1
FT /note= "promoter of E1 gene of rice (W09213956)"
FT 3970..4245
FT *tag= f
FT /product= improved_barstar
FT 4246..4577
FT *tag= g
FT /label= 3'_chs
FT /note= "region containing 3' untranslated end of
FT chalcone synthase gene"
FT complement (5325..5349)
FT *tag= h
FT /note= "T-DNA left border"
FT
XX W09810081-A2.
XX
XX 12-MAR-1998.
XX
XX 01-SBP-1997; 97WO-EP04739.
XX
XX 03-SEP-1996; 96EP-0202446.
XX
XX (PLBZ ) PLANT GENETIC SYSTEMS NV.
XX
XX Michiels F, Williams M;
XX WPI; 1998-193630/17.
XX
XX DNA encoding an improved barstar protein - used to restore fertility
XX in male-sterile plant lines
XX
XX Example 4; Pages 41-43; 54pp; English.
XX
XX The present sequence was used in the preparation of an improved
XX Bacillus amyloliquefaciens barstar, i.e. barnase inhibitor, which
XX can be used to restore fertility to male-sterile lines.
XX The DNA sequence encoding the improved barstar, leads to increased
XX barstar production in tapetum cells, due to improved translation,
XX and possibly protein stability.
XX
XX Sequence 5349 BP; 1339 A; 1233 C; 1290 G; 1487 T; 0 other;

Query Match 15.1%; Score 885.8; DB 19; Length 5349;
Best Local Similarity 99.8%; Pred. No. 1.7e-139;
Matches 887; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 aattacaacggtatatactcgcagtagctcgcctcgaactcgaactcgcctcagtagcatcg 60
Db |||||
QY 1 aattacaacggtatatactcgcagtagctcgcctcgaactcgaactcgcctcagtagcatcg 60
Db |||||
QY 61 gtcgataagaagaaggcaattttagatgtaattcccatcttgaaagaataatagtttaa 120
Db |||||
QY 61 gtcgataagaagaaggcaattttagatgtaattcccatcttgaaagaataatagtttaa 120
Db |||||
QY 121 atatttattgataaaataacagtagcaggtattatagtcgaagcaaaacataaattatt 180
Db |||||
QY 121 atatttattgataaaataacagtagcaggtattatagtcgaagcaaaacataaattatt 180
Db |||||
QY 181 gatcgaagtttaaatcagaataatttcaactgattatcagctgggtacattgcg 240
Db |||||
QY 181 gatcgaagtttaaatcagaataatttcaactgattatcagctgggtacattgcg 240
Db |||||
QY 241 tagatgaagactgagtgcgatattatgtgtaatacataaattgatgatagtagctt 300

```

```

Db 241 tagatgaagactgagtgcgatattatgtgtaatacataaattgatgatagtagctt 300
QY 301 agctcatcggggatcctagaacgctgtagtcagatctcgttgacgggagaccggac 360
Db |||||
QY 301 agctcatcggggatcctagaacgctgtagtcagatctcgttgacgggagaccggac 360
Db |||||
QY 361 ggggcggtaccggcaggtgaagtccagctgccagaaccacacacacacacacacac 420
Db |||||
QY 361 ggggcggtaccggcaggtgaagtccagctgccagaaccacacacacacacacacac 420
Db |||||
QY 421 gcttgaagcggcggccggcggcggcggcggcggcggcggcggcggcggcggcggcgg 480
Db |||||
QY 481 gcaagctcgggttcgttgggagcggcggcggcggcggcggcggcggcggcggcggcgg 540
Db |||||
QY 481 gcaagctcgggttcgttgggagcggcggcggcggcggcggcggcggcggcggcggcgg 540
Db |||||
QY 541 ccagggaactcagcaggtgggtgtagagcgtggagccagtcggcggcggcggcggcgg 600
Db |||||
QY 541 ccagggaactcagcaggtgggtgtagagcgtggagccagtcggcggcggcggcggcgg 600
Db |||||
QY 601 gggagacgtacacggtcgactcggcggcggcggcggcggcggcggcggcggcggcggc 660
Db |||||
QY 601 gggagacgtacacggtcgactcggcggcggcggcggcggcggcggcggcggcggcgg 660
Db |||||
QY 661 ccgctagggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcgg 720
Db |||||
QY 661 ccgctagggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcgg 720
Db |||||
QY 721 ccgcagacggacagaggtcgctccgctccactcctcgggttcctcggctcggtaagg 780
Db |||||
QY 721 ccgcagacggacagaggtcgctccgctccactcctcgggttcctcggctcggtaagg 780
Db |||||
QY 781 tgacctgtcttctcgtatgtagtggttacgatgtgcagacccggcggcggcggcggc 840
Db |||||
QY 781 tgacctgtcttctcgtatgtagtggttacgatgtgcagacccggcggcggcggcggc 840
Db |||||
QY 841 cgggtgcacggcgagtagtcggcggcggcggcggcggcggcggcggcggcggcggc 889
Db |||||
QY 841 cgggtgcacggcgagtagtcggcggcggcggcggcggcggcggcggcggcggcggc 889
Db |||||

RESULT 7
AAZ91097/c
ID AAZ91097 standard; DNA; 6539 BP.
XX
XX AAZ91097;
XX
XX 06-JUN-2000 (first entry)
XX
XX E. coli plasmid pTS431 containing mutant barnase gene.
XX
XX Male sterile plant; mutant barnase gene; anther-specific expression;
XX low fidelity PCR; primer; plant breeding; ss.
XX
XX Synthetic.
XX
XX W0200008176-A1.
XX
XX 17-FEB-2000.
XX
XX 03-AUG-1999; 99WO-JP04167.
XX
XX 04-AUG-1998; 98JP-0220060.
XX
XX (NLSB ) JAPAN TOBACCO INC.
XX
XX Hamada K, Nakakido F;
XX
XX WPI; 2000-195581/17.
XX

```

PT Mutate barnase gene for efficient construction of plant transformants,  
PT particularly male sterile plants free from any undesirable characters  
PS by specifically expressing the gene alone in anther -  
XX Example 3; Page 23-27; 30pp; Japanese.

XX The invention relates to the generation of male sterile plants by  
CC the introduction of a mutant barnase gene (AAZ91095) for expression  
CC specifically in the anther of a plant. This sequence represents the  
CC E. coli/Agrobacterium shuttle vector plasmid pTS172 which contains  
CC the mutated barnase gene (AAZ91095) under control of the cauliflower  
CC mosaic virus 35S promoter. The vector also contains a region of the  
CC Agrobacterium T-DNA gene 7. The vector is used for transmitting the  
CC barnase gene to plants via an Agrobacterium tumefaciens host cell.  
CC The transformed plant is used in plant breeding.

XX Sequence 6539 BP; 1755 A; 1578 C; 1519 G; 1687 T; 0 other;

Query Match 14.2%; Score 832.8; DB 21; Length 6539;  
Best Local Similarity 99.8%; Pred. No. 1.3e-130;  
Matches 834; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 54 gtacatgctgataaagaaagcaattgtagatgttaattcccatcttgaagaaatat 113  
Db 6531 GTACATGCTGATAGAAAGGCAATTGTAGATGTTAATTCCTTGAAGAAATAT 6472  
Qy 114 agtttaaatattattgataaaatacaagtcaggtattatagtcacagcaaaacataa 173  
Db 6471 AGTTTAAATATTATTGATATAAATAACAAGTCAGGTATTATAGTCCAAGCAAAACATAA 6412  
Qy 174 atttattgatgaagtttaattcaagaataatttcaataactgattatcatcagctgtac 233  
Db 6411 ATTTTATTGATGCAAGTTTAAATTCAGAAATATTTCAAATCACTGATTATATCAGCTGTAC 6352  
Qy 234 attccttaataaagacagtgatgatatattatgtataatacaataatgatatag 293  
Db 6351 ATTCCGTAGATGAAGACATGAGTGCCATATATTGTGTAAATACATAAATTGATGATATAG 6292  
Qy 294 ctactgtactcactcggggatcctagaaacgctgtactcagatctcgtggtacgggcagg 353  
Db 6291 CTAGCTTAGCTCATCGGGGATCTCTAGAACGGTGATCTCAGATCTCGGTACGGGCAGG 6232  
Qy 354 accgagcggggcgtaaccgagcgtaagtcacagtcgcagaaacccacatcatgccaag 413  
Db 6231 ACCGAGCGGGCGGTACCGGCGAGCTGAAGTCCAGCTGCCAGAAACCCACATCATGCCAG 6172  
Qy 414 ttcccgacttgagc 473  
Db 6171 TTCCCGTGTGTGAAGCTCG 6112  
Qy 474 tgcctgcgcacgctcgggtgctgtgggcagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 533  
Db 6111 TGCAATGCGCAGGCTCGGTGCTGGGCGAGCCCGATGACAGCGACACCGCTCTTGAAGCCC 6052  
Qy 534 tgtgctccaggggacttcagcaggtggtgtagagcgttgagcccaagtcocgctccgctgg 593  
Db 6051 TGTCCCTCCAGGGACTTTCAGCAGGTGGGTGTAGAGCGTGGAGCCAGTCCCGTCCGCTGG 5992  
Qy 594 tggcgggggagacgtacacagctgactcggcctccagtcgtagcgttcgcttc 653  
Db 5991 TGGCGGGGGAGACGTACACGGTGCACTCGGCCCGCTCAGTGTAGGCGTTCGCTGCCCTTC 5932  
Qy 654 caggggccgcgttaggcgtatgcgggcgacctcgcgtccacctcggcgacgagccagga 713  
Db 5931 CAGGGCCCGGTAGGCGATGTCGGCGACCTCGCGCTCCACCTCGGCACAGCCAGGGA 5872  
Qy 714 tagcgtcccgacagcaggaaggtcgtccogtccactctcgtcgttcgtcggtcgtgta 773  
Db 5871 TAGCGCTCCCGACAGCAGCAGGTGCTGCCGTCCACTCCTCGGTTCTGCGGCTCGGTA 5812  
Qy 774 cggaggtgaccgtgcttgcgtatgtagtggttgacgatggtgcagaccgcgcgcgcgcgcgc 833  
|||||

Db 5811 CGGAAGTTGACCGTGCTGTCTCGATGTAGTGGTTGAGGATGGTCAGACCGCGGCATG 5752  
Qy 834 tccgcctcgggtggcacgcgcatcgccggcgctgcttctggtccattgttct 889  
|||||  
Db 5751 TCCGCTCGGTGGCACGGCGCATCGCGCGGCGCTGCTTCTGGTCCATGGTTAT 5696

RESULT 8

ID AAT39336/c  
XX AAT39336 standard; DNA; 6548 BP.  
AC AAT39336;  
XX 22-JAN-1997 (first entry)  
XX Plasmid pTS174 used to obtain male sterile rice.  
XX Plasmid pTS174; male sterile; barnase; ribonuclease; transgenic plant;  
KW rice; Oryza sativa; ds; cyclic.  
XX Synthetic.  
XX Key Location/Qualifiers  
FT misc\_feature 1..2003 /tag= a  
FT /label= Vector  
FT /note= "pUC19 derived vector sequences"  
FT complement (2019..2283)  
FT /tag= b  
FT /label= 3'nos  
FT /note= "region containing polyadenylation signal  
FT nopaline synthase gene of Agrobacterium  
FT T-DNA"  
FT complement (2284..2624)  
FT /tag= c  
FT /label= Barnase  
FT /product= Bacillus amyloliquefaciens barnase  
FT complement (2625..4313)  
FT /tag= d  
FT /label= PE1  
FT /function= promoter of the stamen-specific E1 gene  
FT Of rice  
FT 4336..5710  
FT /tag= e  
FT /label= P35S  
FT /function= 35S promoter of cauliflower mosaic virus  
FT 5711..6262  
FT /tag= f  
FT /label= bar  
FT /product= phosphinothricin acetyltransferase  
FT 6263..6496  
FT /tag= g  
FT /label= 3'g7  
FT /function= region containing polyadenylation signal  
XX W09626283-A1.  
XX 29-AUG-1996. 96WO-EP00722.  
XX 21-FEB-1996;  
XX 21-FEB-1995; 95EP-0400364.  
XX (PLBZ ) PLANT GENETIC SYSTEMS NV.  
XX Botterman J, Cornelissen M, Michiels F;  
XX WPI; 1996-402373/40.  
XX Prodn. of male sterile plants by transforming with a chimaeric  
PT construct - comprising a male sterility DNA e.g. barnase and a  
PT co-regulating gene, e.g. barstar, into the nuclear genome, useful  
PT for generating hybrid cultivars

QY	834	tcgcctctcggtgcacggcgagatgtcgccggcgctcgcttctgggtccattgttct	888
Db	5760	TCGCGCTCGGTGGCAGCGGGATGTCGCGCGGGCGGTCGTCGGGTCCATGCTTAT	570
RESULT	9		
AAAT61394/C			
ID	AAAT61394	standard; DNA; 6548 BP.	
XX			
AC	AAAT61394;		
XX			
DT	07-MAY-1997	(first entry)	
XX			
DE	Plasmid	PTS172.	
XX			
KW	Transgenic plant; poly-(ADP-ribose) polymerase inhibitor; PARP;		
KW	niacinamide; Agrobacterium; T-DNA; male sterile; barnase;		
KW	ribonuclease; RNase; cereal; wheat; Triticum aestivum;		
KW	plasmid	PTS172; ds.	
XX			
OS	Chimeric Agrobacterium sp.;		
OS	Chimeric Oryza sativa;		
OS	Chimeric cauliflower mosaic virus.		
XX			
FH	Key	Location/Qualifiers	
FT	3'UTR	complement (2019..2288)	
FT		/*tag= a	
FT		/label= 3'nos	
FT		/note= "3' untranslated region contg. the poly-A	
FT		signal of Agrobacterium T-DNA nopaline	
FT		synthase gene"	
FT	CDS	complement (2289..2624)	
FT		/*tag= b	
FT		/product= barnase	
FT	promoter	complement (2625..4313)	
FT		/*tag= c	
FT		/label= PEL	
FT		/note= "promoter region of rice El gene"	
FT	promoter	complement (4336..5710)	
FT		/*tag= d	
FT		/label= P35S	
FT		/note= "35S promoter region of cauliflower mosaic	
FT	CDS	5711..6262	
FT		/*tag= e	
FT		/label= Bar	
FT		/note= "phosphinothricin acetyltransferase"	
FT	3'UTR	6243..6496	
FT		/*tag= f	
FT		/label= 3'g7	
FT		/note= "3' untranslated region contg. the poly-A	
FT		signal of gene 7 of Agrobacterium T-DNA"	
PN	EP757102-A1.		
XX			
XX	05-FEB-1997.		
PD			
XX			
PF	04-AUG-1995;	95EP-0401844.	
XX			
PR	04-AUG-1995;	95EP-0401844.	
XX			
PA	(PLBZ ) PLANT GENETIC SYSTEMS NV.		
PA			
PI	De Block M;		
XX			
DR	WPI; 1997-111050/11.		
XX			
PT	Prodn. of transgenic plants using a poly-(ADP-ribose) polymerase		
PT	inhibitor - reduces the cultured cells response to stress and		
PT	reduces metabolism		
PS	Example 2; Page 17-20; 25pp; English.		
XX			

Plasmid pT5172 (AT61394) contains the barnase coding sequence under control of the rice *El* gene stamen-specific promoter and a phosphinothricin acetyltransferase coding sequence under control of the *CaMV* 35S promoter. Plasmid pT5172 and plasmid pT5772 (see also AT61395) were used to transform wheat Spring variety Pavon calli via particle bombardment. Some calli were treated with the poly-(ADP-ribose) polymerase inhibitor nialamide before, or before and after, bombardment. Healthy, male sterile plants were regenerated only from bombarded calli that were treated with nialamide. This was believed to be due to more faithful expression characteristics of the integrated stamen-selective barnase gene in these calli and regenerated shoots. For plants transformed with pT5172, foreign DNA was stably incorporated in the wheat genome in 2-3 copies.

Sequence 6548 BP; 1756 A; 1579 C; 1523 G; 1690 T; 0 other;

```
Query Match      14.2%; Score 832.8; DB 18; Length 6548;
Best Local Similarity 99.8%; Pred. NO. 1.3e-130;
Matches 834; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

Qy	54	gtacatgggtcgaataaagaaggcaattgtagatgttaattcccatcttgaagaataat	113
Db	6540	GTACATGGTGCATGAAGAAAAGGCAATTTCTAGATGTTAATTCCTCATCTTGAAGAAATAT	6481
Qy	114	agtttaaatatttatgataaaataacaagtcagggtatattagttcccaagcaaaaaacataa	173
Db	6480	AGTTTAAATATTTATTGATAAATAACAAGTCAGGTATTTAGTCCAAAGCAAAACATAA	6421
Qy	174	atttattgatgcgaagttaaatttcagaataattccaataactgattatcatcgctgtac	233
Db	6420	ATTTATTGATGCAAGTTTAAATTCAGAAATATTTCAATACTGATTATTATTCAGCTGGTAC	6361
Qy	234	attgccgtagatgaagactgagtcgatatattgtgtaatacataaaattgatgatag	293
Db	6360	ATTCCCGCTAGATGAAAGACTGAGTGGCATATTATGTGTAATACATAAAATTCATCATATAG	6301
Qy	294	ctaacttagtcatcgggggatactagaacgcgtgatctcagatctcogtgcagggcagg	353
Db	6300	CTAGCTTTAGTCTATCGGGGGATCTTAGAACGGTGTATCTCAGATCTCGGTGCACGGGAGG	6241
Qy	354	accggacggggcggtaccggcaggctgaagtcacagctgccagaaacccacgtcatgccag	413
Db	6240	ACCGGACGGGGCGGTACCGGACGGCTGAAGTCCAGCTGCCAGAAACCCACGCTCATGCCAG	6181
Qy	414	ttcccgctgttgaaagccggccgcgcgcagcatgcccgcggggggggaataatccgagcgctcg	473
Db	6180	TTCCCGTGTCTTGAAGCGGGCGCCGCCAGCATGTCGCCGGGGGATATCCGAGCGCCTCG	6121
Qy	474	tgeatgcacgcgtcgggtcgttgggcagcccgatgacagggaccacgcctcttaagccc	533
Db	6120	TGCATGCGCAGCTTCGGGTGTTTGGGACGCCCGATGACAGCGACCCACGCTCTTGAAGCCC	6061
Qy	534	tgtcctccagggaacttcagcaggtgggtgtagagcgtggagccagtcctcgctcgctgg	593
Db	6060	TGTCCTCCAGGGACTTCACGAGTGGGTGTAGAGGCTGGAGCCCGAGTCCCGCTCGGCTGG	6001
Qy	594	tggcggggggagacgtacacggtcgactcggccgctccagtcgtagggcgttgcgtgccttc	653
Db	6000	TGGCGGGGGAGACGTACACGGTCGACTCGGCCGCTCCAGTCTGTAGCGCTTGCCTGCCTC	5941
Qy	654	cagggggcccggtaggcgatgccggcgcaactcgcgcgtccactcggcgagcagcaggga	713
Db	5940	CAGGGGCCCGGTAGCGGATGCGCGGCGACTCGCCGCTCCACTTCGGGACAGGACGAGGGA	5881
Qy	714	tagcgctcccgcagacgagcagagtcgctccgttcactcctcggttctctcggtcggtta	773
Db	5880	TAGCGCTCCCGCAGACGGACGAGTCTCCGTCCACTCTCTGGCGTTCTCTGGCGCTCGGTA	5821
Qy	774	cggaaagttagccctgcttctctcgatgtagtgtgttgacgatgggtgcagaccgcgcgcatg	833
Db	5820	CGGAAGTTGACCTGCTGTCTCGATGTAGTGGTTTGACATGTTGCAGACGCCCGCGGATG	5761

Qy	834	tccgcctcgtggcacgcggaatcggccggcgtcgtttcttggtccattgtctt	889
Db	5760	TCCGCCCTCGTGGCACGCGGATGTCGGCCGGCGCTCGTTCTGGGTCCATGGTTAT	5701

RESULT 10  
AA291096/C

AAZ91030/C  
ID AAZ91096 standard; DNA; 6548 BP.

XX AA791096:

XX  
DT 06-JUN-2000 (first entry)

DE *E. coli* plasmid pTS172 containing synthetic barnase gene.

XX Male sterile plant; mutant barnase gene; anther-specific expression;  
KW low fidelity PCR; primer; plant breeding; ss.  
KW

YY	Synthetic.
OS	

AA WO200008176-A1. PN

XX  
PD 17-FEB-2000

XX  
PF 03-AUG-1999. 99W0-TP04167

XX  
DB 04-NYC-1000. 98TP-0330060

	2
	3
	4
	5
	6
	7
	8
	9
	10
	11
	12
	13
	14
	15
	16
	17
	18
	19
	20
	21
	22
	23
	24
	25
	26
	27
	28
	29
	30
	31
	32
	33
	34
	35
	36
	37
	38
	39
	40
	41
	42
	43
	44
	45
	46
	47
	48
	49
	50
	51
	52
	53
	54
	55
	56
	57
	58
	59
	60
	61
	62
	63
	64
	65
	66
	67
	68
	69
	70
	71
	72
	73
	74
	75
	76
	77
	78
	79
	80
	81
	82
	83
	84
	85
	86
	87
	88
	89
	90
	91
	92
	93
	94
	95
	96
	97
	98
	99
	100
	101
	102
	103
	104
	105
	106
	107
	108
	109
	110
	111
	112
	113
	114
	115
	116
	117
	118
	119
	120
	121
	122
	123
	124
	125
	126
	127
	128
	129
	130
	131
	132
	133
	134
	135
	136
	137
	138
	139
	140
	141
	142
	143
	144
	145
	146
	147
	148
	149
	150
	151
	152
	153
	154
	155
	156
	157
	158
	159
	160
	161
	162
	163
	164
	165
	166
	167
	168
	169
	170
	171
	172
	173
	174
	175
	176
	177
	178
	179
	180
	181
	182
	183
	184
	185
	186
	187
	188
	189
	190
	191
	192
	193
	194
	195
	196
	197
	198
	199
	200
	201
	202
	203
	204
	205
	206
	207
	208
	209
	210
	211
	212
	213
	214
	215
	216
	217
	218
	219
	220
	221
	222
	223
	224
	225
	226
	227
	228
	229
	230
	231
	232
	233
	234
	235
	236
	237
	238
	239
	240
	241
	242
	243
	244
	245
	246
	247
	248

XX

XX

XX Mutate barnase gene for efficient construction of plant transformants.  
PT Particularly male sterile plants free from any undesirable characters  
PT by specifically expressing the gene alone in another -

Example 3: Page 19-23; 30pp; Japanese.

The invention relates to the generation of male sterile plants by the introduction of a mutant barnase gene (AA951095) for expression specifically in the anther of a plant. This sequence represents the E. coli/Agrobacterium shuttle vector plasmid pTS172 which contains the synthetic barnase gene (AA951094) under control of the cauliflower mosaic virus 35S promoter. The vector also contains a region of the Agrobacterium T-DNA gene 7. The vector is used for transmitting the barnase gene to plants via an Agrobacterium tumefaciens host cell. The transformed plant is used in plant breeding.

Sequence 6548 BP: 1756 A: 1579 C: 1523 G: 1690 T: 0 other: XX  
SO

Query Match	Score	DB	Length
Best Local Similarity	99.8%	21	6548
Matches 834: Conservative	Pred. NO. 1.3e-130		
	0: Mismatches	2: Indels	Gaps

QY 54 gtacatggtcgataagaaaaaggcaattttagatgttaattcccatcttgaagaatat 113  
 |||||  
 pb 6540 GTACATGGTCGATAGAAAAAGSCAAATTGTAGATGTTAATCCCATCTTGAAGAATAAT 6481

QY 114 agtttaaattatttattgataaaataaacaagtcagggtatttatagtcgaagcaaaaaacataa 173  
 |||||  
 db 6480 AGTAAATATTATTATGATAAAATTAACAAGTCAGGTATTATAGTCCAAAGCAAAACATAA 6421

QY 174 atttattgatgcaagttaaattcaagaataatttcaataactgattatatcagctggtac 233  
 |||||  
 db 6420 attttattgatgcaagttttaaattcagaataatttcaataactgattatatcagctggtac 6361

QY 234 attgccgtagatgaagactgagtcgcatattatgtgtaatacacataaattgatgatatag 293  
|||||  
db 6360 attgccgtagatgaagactgagtcgcatattatgtgtaaacacataaattgatgatatag 6301



Qy 826 ccggcatgcgcctcggtggcagcgcgatgtcgccggcgctcgcttctgggtccattg 885  
 |||||  
 Db 5988 CCGGCATGTCGGCTCGGTGCGACGGCGATGTCGGCGGCGTCTTCTGGGCTCATGG 5929  
 Qy 886 ttcttcttt 894  
 |||||  
 Db 5928 ATCCGATTT 5920

RESULT 12  
 AAT39338  
 ID AAT39338 standard; DNA; 3658 BP.  
 XX  
 AC AAT39338;  
 DT 22-JAN-1997 (first entry)  
 XX  
 DE Plasmid pVE136 (EcoRI-HindIII fragment).  
 XX  
 KW Plasmid pVE136; male sterile; barnase; ribonuclease; barstar;  
 transgenic plant; maize; corn; Zea mays; ds.  
 XX  
 OS Synthetic.  
 XX

Key Location/Qualifiers  
 misc\_feature 1..26  
 FT /\*tag= a  
 FT /label= pUC19  
 FT /note= "polylinker of pUC19"  
 FT complement (28..403)  
 FT /\*tag= b  
 FT /label= 3'nos  
 FT /note= "region containing polyA signal of  
 Agrobacterium T-DNA nopaline synthase gene"  
 CDS complement (404..739)  
 FT /\*tag= c  
 FT /label= Barnase  
 FT /note= "Bacillus amyloliquefaciens barnase coding  
 region"  
 FT complement (740..1918)  
 FT /\*tag= d  
 FT /label= PCA55  
 FT /note= "maize CA55 gene promoter"  
 1956..2788  
 FT /\*tag= e  
 FT /label= p35S  
 FT /note= "cauliflower mosaic virus 35S promoter"  
 2789..3340  
 FT /\*tag= f  
 FT /label= bar  
 FT /note= "phosphinothricin acetyltransferase coding  
 region"  
 FT 3341..3623  
 FT /\*tag= g  
 FT /label= 3'nos  
 FT /note= "region containing polyA signal of  
 Agrobacterium T-DNA nopaline synthase gene"  
 3624..3658  
 FT /\*tag= h  
 FT /label= pUC19  
 FT /note= "polylinker of pUC19"

W09626283-A1.  
 XX  
 XX 29-AUG-1996.  
 XX  
 XX 21-FEB-1996; 96WO-EP00722.  
 XX  
 XX 21-FEB-1995; 95EP-0400364.  
 XX (PLBZ ) PLANT GENETIC SYSTEMS NV.  
 XX  
 PI Botterman J, Cornelissen M, Michiels F;

XX WPI; 1996-402373/40.  
 DR  
 XX  
 PT Prodn. of male sterile plants by transforming with a chimaeric  
 construct comprising a male sterility DNA e.g. barnase and a  
 PT co-regulating gene, e.g. barstar, into the nuclear genome, useful  
 PT for generating hybrid cultivars  
 XX  
 PS Example 2; Page 40-42; 56pp; English.  
 XX  
 CC The EcoRI-HindIII fragment (AAT39338) of plasmid pVE136 contains  
 CC a Bacillus barnase (ribonuclease) gene under control of the PCA55  
 CC promoter. Embryogenic callus from maize was transformed with  
 CC pVE136 and pTS88 (see also AAT39337), a pGEM2-derived plasmid contg.  
 CC barstar DNA under control of a 35S promoter. Maize plants contg.  
 CC both p35S-barstar-3'g7 and the PCA55-barnase-3'nos chimeric genes  
 CC were male sterile. Foreign genes such as barnase can be placed  
 CC under control of a stamen-specific promoter to produce male sterile  
 CC plants; constitutive expression of barstar counteracts any low  
 CC level expression of barnase in other tissues.  
 XX  
 SQ Sequence 3658 BP; 1050 A; 857 C; 797 G; 954 T; 0 other;

Query Match 12.6%; Score 741.8; DB 17; Length 3658;  
 Best Local Similarity 99.7%; Pred. No. 1.9e-115;  
 Matches 743; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2630 gaattcgagctcggtaccgggggatcttcccgatcttagtaacatagatgacacgcgcgc 2689  
 |||||  
 Db 1 gaattcgagctcggtaccgggggatcttcccgatcttagtaacatagatgacacgcgcgc 60  
 Qy 2690 gataatttacctagttgcgcgtatatattgtttctatcgcgtattaaatgtataat 2749  
 |||||  
 Db 61 gataatttacctagttgcgcgtatatattgtttctatcgcgtattaaatgtataat 120  
 Qy 2750 tgcgggactctaatcataaaaaaccatctctataataacgtcatgtacattgttaatt 2809  
 |||||  
 Db 121 tgcgggactctaatcataaaaaaccatctctataataacgtcatgtacattgttaatt 180  
 Qy 2810 attacatgcttaacgtaattcaacagaaattatatgataatcgcgaagaccggcaaca 2869  
 |||||  
 Db 181 attacatgcttaacgtaattcaacagaaattatatgataatcgcgaagaccggcaaca 240  
 Qy 2870 ggattcaatcttaagaaactttattgccaaatgtttgaacgatctgctcgatcctcta 2929  
 |||||  
 Db 241 ggattcaatcttaagaaactttattgccaaatgtttgaacgatctgctcgatcctcta 300  
 Qy 2930 gagcggaaagtgaattgaccgatcagagtttgaagaaaaattattacacactttatg 2989  
 |||||  
 Db 301 gagcggaaagtgaattgaccgatcagagtttgaagaaaaattattacacactttatg 360  
 Qy 2990 taaagctgaaaaaacggcctcccgaggaaacggcttttttcgttatctgattttgtaa 3049  
 |||||  
 Db 361 taaagctgaaaaaacggcctcccgaggaaacggcttttttcgttatctgattttgtaa 420  
 Qy 3050 aggtctgataatggtccggtttttttaaatacagccagctcgttgagtaaagaatccggt 3109  
 |||||  
 Db 421 aggtctgataatggtccggtttttttaaatacagccagctcgttgagtaaagaatccggt 480  
 Qy 3110 ctgaatttctgaagcctgatgtatagtttaataatccggttcacgcgatgttcgctcgcttt 3169  
 |||||  
 Db 481 ctgaatttctgaagcctgatgtatagtttaataatccggttcacgcgatgttcgctcgcttt 540  
 Qy 3170 tgcggggagtttgcccttccctgtttgagaagatgtctcccgccgatgttttcccgagg 3229  
 |||||  
 Db 541 tgcggggagtttgcccttccctgtttgagaagatgtctcccgccgatgttttcccgagg 600  
 Qy 3230 cgacgtctgcaaggttcccttcttctggtccaccagcgaggttgcttctctgattttg 3289  
 |||||  
 Db 601 cgacgtctgcaaggttcccttcttctggtccaccagcgaggttgcttctctgattttg 660  
 Qy 3290 taatgtaattatcaggttagcttatgatattgtgaagataatccgcaaaccccgctcaaacg 3349

```

Db 661 taatgaataacaggtagcttatgatctgtaagataatccgcaacccggtcaaacg 720
Qy 3350 tgttgataacccggtaccattggttagc 3374
Db 721 tgttgataacccggtaccattggttagc 745

RESULT 13
AAQ53889
ID AAQ53889 standard; DNA; 4808 BP.
XX
AC AAQ53889;
XX
DT 27-JUN-1994 (first entry)
XX
DE Restriction fragment of construct carrying plant maintainer gene.
XX
KW Maintainer gene; sterile; sterility; homogenous population; hybrid;
KW seed; fertility restorer gene; pollen lethality gene; ss.
OS Synthetic.
XX
FH Key
FT misc_signal
FT complement (18..401)
FT /tag= a
FT /label= 3', nos.
FT /note= "3', regulatory sequence containing the
FT polyadenylation site derived from
FT Agrobacterium T-DNA nopaline synthase gene."
FT
FT CDS
FT complement (402..737)
FT /tag= b
FT /label= Barnase.
FT /note= "Coding region of the barnase gene of
FT Bacillus amyloliquefaciens."
FT
FT promoter
FT complement (738..1944)
FT /tag= c
FT /label= PZM13.
FT /note= "Promoter region of the Zm 13 gene of Zea
FT mays."
FT
FT misc_signal
FT complement (1945..2281)
FT /tag= d
FT /label= 3', nos.
FT
FT CDS
FT complement (2282..2554)
FT /tag= e
FT /label= Barstar.
FT /note= "Coding region of the barstar gene of
FT Bacillus amyloliquefaciens."
FT
FT promoter
FT complement (2555..3099)
FT /tag= f
FT /label= PTA29.
FT /note= "Promoter region of the TA29 gene of
FT Nicotiana tabacum."
FT
FT promoter
FT 3100..3932
FT /tag= g
FT /label= 35S3.
FT /note= "35S3" promoter sequence derived from
FT cauliflower mosaic virus isolate CabbB-JI."
FT
FT CDS
FT 3933..4484
FT /tag= h
FT /note= "Coding region of the phosphinothricin
FT acetyltransferase gene."
FT
FT misc_signal
FT 4485..4763
FT /tag= i
FT /label= 3', nos.
XX
PN W09325695-A.
XX
PD 23-DEC-1993.
XX
PF 11-JUN-1993; 93WO-EP01489.
XX
PR 12-JUN-1992; 92US-0899072.

```

```

PR 03-NOV-1992; 92US-0970840.
XX
PA (PLBZ ) PLANT GENETIC SYSTEMS NV.
XX
PI Leemans J, Williams M;
XX
DR WPI; 1994-007552/01.
XX
PT Maintainer gene for maintenance of male-sterile plants -
PT comprises fertility-restorer gene and pollen-lethality gene
XX
PS Example 2; Page 73-75; 87pp; English.
XX
CC A maintainer gene of plants, pref. a foreign chimeric gene,
CC comprises (a) a fertility restorer gene which comprises a fertility
CC restorer DNA and (ii) a restorer promoter capable of
CC directing the expression of the fertility restorer DNA and (b) a
CC pollen lethality gene that is selectively expressed in microspores
CC and/or pollen of the plant to prevent the production of functional
CC pollen and which comprises (i) a pollen lethality DNA and (ii) a
CC pollen specific promoter capable of directing expression of the
CC pollen lethality DNA. Plants transformed with this DNA (maintainer
CC plants) can be used to maintain a homogenous population of male
CC sterile plants for the production of hybrid seed. This sequence is
CC an EcoRI-HindIII restriction fragment of the construct designated
CC pT5218 and comprises the maintainer gene described.
XX
SQ Sequence 4808 BP; 1370 A; 1064 C; 1037 G; 1333 T; 4 other;

Query Match 12.4%; Score 727.8; DB 15; Length 4808;
Best Local Similarity 99.2%; Pred. NO. 4.4e-113;
Matches 743; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

Qy 2630 gaattcgagctcggtaccggggtatctcccgatctccgtagtaacatagatgacacccgcgcc 2689
Db 1 gaattcgagctcggtaccggggtatctcccgatctccgtagtaacatagatgacacccgcgcc 60
Qy 2690 gataattatctcgtggtgctgctatattgttttctatcggtattataatgataat 2749
Db 61 gataattatctcgtggtgctgctatattgttttctatcggtattataatgataat 120
Qy 2750 tgcgggactctaatcataaaaaacccatctcctataataacgctcatcattacatgttaatt 2809
Db 121 tgcgggactctaatcataaaaaacccatctcctataataacgctcatcattacatgttaatt 180
Qy 2810 attacatgcttaacgtaattcaacagaaattatatgataatcgcgaagccggcaaca 2869
Db 181 attacatgcttaacgtaattcaacagaaattatatgataatcgcgaagccggcaaca 240
Qy 2870 ggattcaattttaagaacctttattgccaattgttgccaatgtttgacgatctgctcgatcctcta 2939
Db 241 ggattcaattttaagaacctttattgccaatgtttgacgatctgctcgatcctcta 300
Qy 2930 gag-----ccggaagtgaaattgaccgatcagagtttgaagaaaaattattacacatt 2985
Db 301 gagnnnnccggaagtgaaattgaccgatcagagtttgaagaaaaattattacacatt 360
Qy 2986 tatgtaagctgaaaaaacgcctcccgaggaagcggcttttctgattctgattttt 3045
Db 361 tatgtaagctgaaaaaacgcctcccgaggaagcggcttttctgattctgattttt 420
Qy 3046 gtaaggctgataatggtccgtgttttgaatacagccagtcgcttgagtaagaatc 3105
Db 421 gtaaggctgataatggtccgtgttttgaatacagccagtcgcttgagtaagaatc 480
Qy 3106 cggctgaattctgaagcctgtagtattgataatccgctcagccatggttcgcgcg 3165
Db 481 cggctgaattctgaagcctgtagtattgataatccgctcagccatggttcgcgcg 540
Qy 3166 cttttgcccgggagtttgcttccctgtttgagaagatgtctccgcgcgatttttcccc 3225
Db 541 cttttgcccgggagtttgcttccctgtttgagaagatgtctccgcgcgatttttcccc 600

```



Qy 3226 ggagcgactgtgcaagggtcccttttgatgccaccagccagcgagggttggtcttgat 3285  
 Db 601 ggagcgactgtgcaagggtcccttttgatgccaccagccagcgagggttggtcttgat 660  
 Qy 3286 ttgttaatgtaattatcaggtagcttatgatctgtcgaagataatccgcaaccccgatca 3345  
 Db 661 ttgttaatgtaattatcaggtagcttatgatctgtcgaagataatccgcaaccccgatca 720  
 Qy 3346 aacgtgttgataaccggtaccatgtagc 3374  
 Db 721 aacgtgttgataaccggtaccatgtagc 749  
 RESULT 14  
 AAT08976/C  
 ID AAT08976 standard; DNA; 4896 BP.  
 XX  
 AC AAT08976;  
 XX  
 DT 12-AUG-1996 (first entry)  
 XX  
 DE pTS256 contg. p35S-bar-3'nos and pTA29-barstar-3'nos.  
 XX  
 KW Cl; Cl-S; transgenic plant; male sterility; colour-linked restorer;  
 KW anthocyanin; aleurone; B-peru; bar; barstar; bar\*; ds.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT misc\_signal complement (39..317)  
 FT /label= a  
 FT /note= "3' regulatory sequence contg. the  
 FT polyadenylation signal of the nopaline  
 FT synthase gene of Agrobacterium T-DNA"  
 FT CDS complement (318..869)  
 FT /label= b  
 FT /note= "coding region of bar gene of  
 FT Streptomyces hygroscopicus"  
 FT promoter complement (870..1702)  
 FT /label= c  
 FT /note= "35S promoter of Cauliflower Mosaic Virus"  
 FT complement (1740..2284)  
 FT /label= d  
 FT /note= "promoter of TA29 gene of Nicotiana  
 FT tabacum"  
 FT CDS 2285..2557  
 FT /label= e  
 FT /note= "coding region of barstar gene of  
 FT Bacillus amyloliquefaciens"  
 FT misc\_signal 2558..2879  
 FT /label= f  
 FT /note= "3' regulatory sequence contg. the  
 FT polyadenylation signal of the nopaline  
 FT synthase gene of Agrobacterium T-DNA"  
 FT misc\_RNA 1..38  
 FT /label= g  
 FT /note= "pUC19  
 FT 2880..4986  
 FT /label= h  
 FT /note= "pUC19 derived sequence"  
 FT misc\_RNA  
 FT /label= pUC19  
 FT /note= "pUC19 derived sequence"  
 XX  
 PN WO9534634-A2.  
 XX 21-DEC-1995.

XX 06-JUN-1995; 95WO-EP02157.  
 XX 06-JUN-1994; 94US-0254776.  
 XX (PUBZ ) PLANT GENETIC SYSTEMS NV.  
 XX Kriebbers E, Leemans J, Williams M;  
 XX WPI; 1996-049664/05.  
 XX Transgenic plants contg. male sterility and colour-linked restorer  
 PT genotypes - used for prodn. of male sterile seeds identifiable from  
 PT their colour, also new truncated anthocyanin regulatory genes and  
 PT aleurone specific promoters  
 XX  
 PS Disclosure; Page 68-71; 104pp; English.  
 XX  
 CC Shortened forms of anthocyanin (An) regulatory genes (ARG) are easier  
 CC to manipulate than the complete gene and still provide An prodn.  
 CC The Cl gene (and the Cl-S gene) can be considerably shortened  
 CC while still retaining, under appropriate conditions, its  
 CC capability of conditioning anthocyanin prodn. in the aleurone  
 CC of seeds of cereal plants such as corn. A pref. shortened Cl gene  
 CC is comprised in pCOL9 (see AAT08975). The full Cl gene sequence is  
 CC given in AAT08973.  
 CC A truncated B-peru gene (presumed sequence = AAT08674; actual  
 CC sequence = AAT08977) may also be used.  
 CC Plasmids pTS256 (AAT08976) and pTS200 (AAT17246) were used in the  
 CC construction of vectors comprising the Cl and B-peru genes  
 CC as well as male-sterility gene and a selectable marker gene.  
 XX  
 SQ Sequence 4896 BP; 1252 A; 1146 C; 1186 G; 1312 T; 0 other;  
 Query Match 12.1%; Score 712; DB 17; Length 4896;  
 Best Local Similarity 76.7%; Pred. No. 1.9e-110;  
 Matches 987; Conservative 0; Mismatches 185; Indels 114; Gaps 5;  
 Qy 2628 cggattcgactcggtaccggggtctcccgatctctagtaacatagatgacaccgcgc 2687  
 Db 2911 CAGAAATTCGAGCTCGGTACCGGGGATCTCCCGATCTAGTAACATAGATGACACCGGC 2852  
 Qy 2688 gcgataattatcttagttgctgcgctatatattgttttctatcgctgattataatgata 2747  
 Db 2851 GCGATAATTTATCTAGTTGCGGCTATATTTGTTTCTATCGCGTATTAAATGTATA 2792  
 Qy 2748 attcgggactctaatcataaaacccatctcataataaacgtcatgattacatgttaa 2807  
 Db 2791 ATTGCGGACTCTAATCATATAAAACCCATCTCATATAAAGTCATGCTACATGTATA 2732  
 Qy 2808 ttattacatgcttaacgttaattcaacagaaattatgataatcgcgaagcgcgcaa 2867  
 Db 2731 TTATTACATGCTTAACGTAATTCAACAGAAATTTATGATGATTAATCATCGCAGACCGGCA 2672  
 Qy 2868 caggattcaatttaagaacctttattgccaaattgtttgaacgactctgctcgatctc 2927  
 Db 2671 CAGGATTCAATCTTAAGAAACTTTATTGCCAAATGTTTGAACGATCTGCTTCGGATCCTC 2612  
 Qy 2928 tagagccggaagtgaaattgaccgcatcagagtttgagaagaaaaatttattacacactta 2987  
 Db 2611 TAGACCAAGCTTCGCGGTTGTGTTTCCATATT-----GTTCAATCTCCCATGTAT 2562  
 Qy 2988 tgtaagctgtaaaaaaacgcctccgcgaggaagcgcgttttttcgtatctatctgtttgt 3047  
 Db 2561 CGTATTAAAGAAAGTATGATGGTGATGTCGCAGCCTTCGCGTTCGCTTCACGGAAACCT 2502  
 Qy 3048 aaaggtctgataatggtccgtttgttttgataaacagccagctcgcttgagtaaaagaatccg 3107  
 Db 2501 GAAGCACACTCTCGGCGCCATT-----TTCACTCAGCTGCTGCTTTGTTTCAAACT 2451  
 Qy 3108 gctcgaattctcgaagcctgtagttagttaaataatccgcttcacgcgaatgttcgtccgcct 3167

Db	2450	GCCTCCATTCCAAACAGAGCGGGTACTCCACCCATCCGGT-----	2409
Qy	3168	ttfgccggggagtttgctctccctggttgagaagatgtccgcgcgagtgctttcccccg	3227
Db	2410	-----CAGACAAATCCCATAAAGCGTCAGGTTTCCACGTAAGTATCCGGAAG	2363
Qy	3228	agcgagctctgcaaggttcctctttgatgccaccacgcgaggctgtgctctcgattt	3287
Db	2362	GCG-----AAGCTCCCTTTTCAATGCTGTGTGAGGTGCGTGATACTACTGTGATTT	2313
Qy	3288	tgtaatgtaattacccggtagettaTgatgtctgaagaTaatccgcaaccgcgcTcaa	3347
Db	2312	-----GTTCCCC	2306
Qy	3348	cggttgatacccggtaccatggtagtagtaattctttaagtaaaactttgatttgagtg	3407
Db	2305	GTTAATCAGTCCTTTTTCATGTAAGTAATTCCTTTAAGTAAATACTTTGATTTCAGTG	2246
Qy	3408	atgagtgttgtactgttcacacttgcaccacgaagggcataataagagcaacagacataca	3467
Db	2245	ATGATGTGTGACTGTTACACTTGCACCACAAAGGCATATATAGAGCAACAAGATACACA	2186
Qy	3468	acaacttgcaaaactaacttttgttgagacatttcggaggaataatggggagtagcagctta	3527
Db	2185	ACAACTTGCAAAACATAACTTTTGTGTGAGCATTTTCGAGAAATATGGGAGTAGCAGGCTA	2126
Qy	3528	atctgagggTcaacattaaaggtttcatgtattaaattgttcgaacacatggacttagtgta	3587
Db	2125	ATCTGAGGGTTAACTAATTAAGTTTTCATGTATTAATTTGTTGCAACATGGACTTAGTGTGA	2066
Qy	3588	ggaaaagTaccaaaattttgtctcaacctgattctcagttatgTgaattacattatgaag	3647
Db	2065	GGAAAAGTACCAAAATTTTGTCTCACCTGATTTTCAGTTATGGAATTTACATTAATGAAG	2006
Qy	3648	ctgtctagagaagatgtttattcttagtccagccaccaccttatgcaagtcgtcttta	3707
Db	2005	CTGTGCTAGAGAAGATGTTTATTCTGTGACGACCCACCTTATGCAAGTCTGCTTTTA	1946
Qy	3708	gcttgattcaaaaactgatttaatttacattgTcaaatgtgcatacttcgagccatgtc	3767
Db	1945	GCTTGATTCAAAAACGTGATTTAAATTTACATTCCTAAATGTGCATCTTCGAGCCATATGTC	1886
Qy	3768	gcttaactcggtagtagtgatatatttagcacataaaaaaatcatgtttgaaatcatcttt	3827
Db	1885	GCTTTAATTCGAGTAGATGCTATATATTAGTACATAAAAAATCATGTTTGAATCATCTTT	1826
Qy	3828	cataagtgcaeaagTcaattgtccctctctgtttggcacataatcaatctgttaatgca	3887
Db	1825	CATAAAGTGCAAGTCAATTTGTCCTTCCTTGTGTTGGCACTATATTCAATCTGTTAAATGCA	1766
Qy	3888	aattatccagttataacttagctagat	3913
Db	1765	AATTATCCAGTTATACTTAGCTAGAT	1740

RESULT 15  
AA15631/C  
ID AA15631 standard; DNA; 4896 BP.

AC AAX15631;

XX  
DT 07-MAY-1999 (first entry)

DE Nucleotide sequence of plasmid pTS256.

C1 gene: maize; male-sterile corn line; anthocyanin production; pTS256; ds.

xx  
OS  
Synthetic.

XX  
DN

XX  
000000000000

PD 09-MAR-1999.

XX	
PF	07-JUN-1995; 95US-0485139.
XX	
PR	07-JUN-1995; 95US-0485139.
PR	06-JUN-1994; 94US-0254776.
XX	
PA	(PLBZ ) PLANT GENETIC SYSTEMS NV.
XX	
PI	Krebberts E, Leemans J, Williams M;
XX	
DR	WPI; 1999-204053/17.
XX	
PT	Process for maintaining a male-sterile corn line - using male
PT	sterile plants lacking functional regulatory gene for anthocyanin
XX	production
XX	
PS	Example 2; Columns 31-36; 35pp; English.
XX	
CC	The present sequence represents the nucleotide sequence of plasmid
CC	pt5256. It is used in the course of the invention. The specification
CC	describes a process for maintaining a male-sterile corn line, using male
CC	sterile parent plants lacking a functional gene for anthocyanin
CC	production, and a maintainer corn line comprising male fertile parent
CC	plants containing foreign DNA comprising a restorer gene and an active
CC	regulatory protein gene. By using the anthocyanin gene, the colour of
CC	the male-sterile plants will differ from that of the male fertile
CC	plants. This will enable the seeds harvested from the plants to be
CC	easily separated into those that will grow into male-fertile plants and
CC	those that will grow into male-sterile plants.
XX	
SQ	Sequence 4896 BP; 1251 A; 1147 C; 1186 G; 1312 T; 0 other;

Query Match 12.1%; Score 708.8; DB 20; Length 4896;  
Best Local Similarity 76.6%; Pred. No. 6.6e-110;  
Matches 985; Conservative 0; Mismatches 187; Indels 114; Gaps 5;

**Qy** 2628 cggaattcgaqctcggtacccgggagatcttcccgatctactaacctcaatcatcttcac

Db

Qy 2688 gcgataattatcctagtttgcgcgctatatattttgtttctatcgcgatattaagtata 2747

DB 2852 GCGATAAATTTATCCTAGTTTGC GCGCTAATTTGTTCTATCGGTATTAATGATA 2793

Db 2792 ATTGGGGACTCTAATCATAAAAACCCATCTCATAAATAACGTCATGCATTACATGTTAA 2733

Qy	2808	2867
	ttattacatgcttaacgtaattcaacagagaattatatgataatcatcgcaagaccggcaa	

Db 2732 TTATTACATGCTTAACGTAATTCACAGAAATTATGATAATCATCGCAAGACGGCAA 2673

2671		CAGGATCCAACTTAAAGAACTTTATGCCAAATGTTTGAACGATCTGCTTGGATCCCTC	2676
2672		CAGGATCCAACTTAAAGAACTTTATGCCAAATGTTTGAACGATCTGCTTGGATCCCTC	2677
Db		CAGGATCCAACTTAAAGAACTTTATGCCAAATGTTTGAACGATCTGCTTGGATCCCTC	2682

Qy 2928 tagagccggaagtgaattgaccgatacagagtttgaagaaaaattttattacacacttta 2987

[illegible][illegible]

3048 aaaggctctgataatgggtccggtgttttgataatcagccagtcgcttgagtaaaagaatccg 3107

2502 GAAGCACACTCTCGGGCGCATT-----TTCAGTCAGCTGCTTGCCTTGTGTTCAAACT 2452

**DY** 3108 gctcgaatttctgaaagcgtgatgtatagttaataatccgttcacgccattcgctcgctcgcgt 3167  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
**bB** 2451 GCCTCCATTCCAAAAGCAGCGGGTACTCCACCAATGCGCM-  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

----- 2410

Search completed: November 15, 2001, 20:34:22  
Job time: 15044 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2001, 16:22:28 ; Search time 154.69 Seconds  
(without alignments)  
8586.809 Million cell updates/sec

Title: US-09-698-903B-1  
Perfect score: 5865  
Sequence: 1 aattacaacggtatatcc.....caattgaatatctgcgcg 5865

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5843.4	99.6	5864	3	US-08-894-440-4
2	4903.2	83.6	4946	3	US-08-817-188-1
3	2386.8	40.7	5560	3	US-08-817-188-5
4	1270.4	21.7	5620	1	US-08-104-072B-9
5	1270.4	21.7	5620	1	US-08-351-413-3
6	1270.4	21.7	5620	2	US-09-025-583-3
7	1269	21.6	1287	1	US-08-064-121-3
8	1269	21.6	1287	1	US-08-478-015-3
9	1269	21.6	1287	3	US-08-475-975-3
10	1269	21.6	1287	3	US-09-084-889-3
11	832.8	14.2	6548	3	US-08-817-188-2
12	831.2	14.2	6548	3	US-08-894-440-1
13	741.8	12.6	3658	3	US-08-894-440-3
14	729.4	12.4	4808	1	US-08-351-413-17
15	729.4	12.4	4808	2	US-09-025-583-17
16	728.2	12.4	6291	3	US-08-817-188-4
17	712	12.1	4896	3	US-08-750-357-2
18	708.8	12.1	4896	2	US-08-485-139-2
19	705.4	12.0	6376	1	US-08-104-073-5
20	705.2	12.0	6555	1	US-08-351-413-2
21	705.2	12.0	6555	2	US-09-025-583-2
22	595.2	10.1	1303	3	US-08-894-440-2
23	564	9.6	790	4	US-09-363-970-4
24	561.6	9.6	2345	3	US-09-026-673-1
25	561.6	9.6	3544	2	US-08-485-139-3
26	561.6	9.6	3544	3	US-08-750-357-3
27	561.6	9.6	3658	3	US-08-894-440-3

c 28	561.6	9.6	4808	1	US-08-351-413-17	Sequence 17, Appl
c 29	561.6	9.6	4808	2	US-09-025-583-17	Sequence 17, Appl
c 30	561.6	9.6	4883	1	US-08-064-121-4	Sequence 4, Appl
c 31	561.6	9.6	4883	1	US-08-318-772A-1	Sequence 1, Appl
c 32	561.6	9.6	4883	1	US-08-478-015-4	Sequence 4, Appl
c 33	561.6	9.6	4883	3	US-08-475-975-4	Sequence 4, Appl
c 34	561.6	9.6	4883	3	US-09-084-889-4	Sequence 4, Appl
c 35	561.6	9.6	4896	2	US-08-485-139-2	Sequence 2, Appl
c 36	561.6	9.6	4896	3	US-08-750-357-2	Sequence 2, Appl
c 37	561.6	9.6	5620	1	US-08-104-072B-9	Sequence 9, Appl
c 38	561.6	9.6	5620	1	US-08-351-413-3	Sequence 3, Appl
c 39	561.6	9.6	5620	2	US-09-025-583-3	Sequence 3, Appl
c 40	561.6	9.6	6291	3	US-08-817-188-4	Sequence 4, Appl
c 41	561.6	9.6	6376	1	US-08-104-073-5	Sequence 5, Appl
c 42	549	9.4	549	5	PCT-US91-0623A-1	Sequence 1, Appl
c 43	544.8	9.3	549	1	US-08-361-467B-10	Sequence 10, Appl
c 44	544.8	9.3	549	1	US-08-484-332C-10	Sequence 10, Appl
c 45	491.4	8.4	6548	3	US-08-894-440-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1  
US-08-894-440-4  
; Sequence 4, Application US/08894440  
; Patent No. 6025546  
; GENERAL INFORMATION:  
; APPLICANT: PLANT GENETIC SYSTEMS N.V.  
; TITLE OF INVENTION: Method to obtain male sterile plants  
; FILE REFERENCE: NMSCOR  
; CURRENT APPLICATION NUMBER: US/08/894,440  
; CURRENT FILING DATE: 1997-11-12  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 5864  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: T-DNA of  
; OTHER INFORMATION: plasmid pTCO113  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: Complement((1)..(25))  
; OTHER INFORMATION: right border of Agrobacterium T-DNA (RB)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: Complement((98)..(330))  
; OTHER INFORMATION: region containing polyadenylation signal of gene 7  
; OTHER INFORMATION: of Agrobacterium T-DNA (3'g7)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: Complement((331)..(882))  
; OTHER INFORMATION: region coding for phosphinothricin acetyl  
; OTHER INFORMATION: transferase (bar)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: Complement((883)..(2608))  
; OTHER INFORMATION: promoter of small subunit gene of Rubisco of  
; OTHER INFORMATION: Arabidopsis (Pssu)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: Complement((2659)..(3031))  
; OTHER INFORMATION: region containing polyadenylation signal of  
; OTHER INFORMATION: nopaline synthase gene of Agrobacterium T-DNA  
; OTHER INFORMATION: (3'nos)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: Complement((3032)..(3367))  
; OTHER INFORMATION: region coding for barnase of Bacillus  
; OTHER INFORMATION: amyloliquefaciens  
; FEATURE:

```
/ NAME/KEY: misc_feature
/ LOCATION: Complement((3368)..(4877))
/ OTHER INFORMATION: promoter of stamen-specific TA29 gene of Nicotiana
/ OTHER INFORMATION: tabacum (PTA29)
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (4924)..(5216)
/ OTHER INFORMATION: promoter of nopaline synthase gene of
/ OTHER INFORMATION: Agrobacterium T-DNA (Phos)
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (5217)..(5489)
/ OTHER INFORMATION: region coding for barstar of Bacillus
/ OTHER INFORMATION: amyloliquefaciens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (5490)..(5765)
/ OTHER INFORMATION: region containing polyadenylation signal of gene 7
/ OTHER INFORMATION: of Agrobacterium T-DNA (3'g)
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: Complement((5840)..(5864))
/ OTHER INFORMATION: left border of Agrobacterium T-DNA
/ US-08-894-440-4
```

```
Query Match          99.6%; Score 5843.4; DB 3; Length 5864;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 5858; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
```

```
Qy 1 aattacaacggtatatactcctgcagctactcggcgctgaactcggcgctcgagtaacatg 60
Db 1 aattacaacggtatatactcctgcagctactcggcgctgaactcggcgctcgagtaacatg 60
Qy 61 gtcgataaagaaaggcaattgttagatgttaattcccatcttgaagaataatagtttaa 120
Db 61 gtcgataaagaaaggcaattgttagatgttaattcccatcttgaagaataatagtttaa 120
Qy 121 atattttatgataaaataacaagtcaggtattatagttccatagtcgaagcaaaacataattatt 180
Db 121 atattttatgataaaataacaagtcaggtattatagttccatagtcgaagcaaaacataattatt 180
Qy 181 gatgcaagtttaattcagaataatttccaataactgattatatacagctggtacattgcgg 240
Db 181 gatgcaagtttaattcagaataatttccaataactgattatatacagctggtacattgcgg 240
Qy 241 tagatgaagactgagtcgcatattatgtgtaatacataaaattgatagtagctt 300
Db 241 tagatgaagactgagtcgcatattatgtgtaatacataaaattgatagtagctt 300
Qy 301 agctcatcggggatcctagaacgcgtgatctcagatctcgtgacgaaacccacgtcatgccgttccgt 360
Db 301 agctcatcggggatcctag-acgcgtgagatcagatctcgtgacgaaacccacgtcatgccgttccgt 360
Qy 361 gggcggtacccgacggctgaagtcacagtcgccagaaacccacgtcatgccgttccgt 420
Db 361 gggcggtacccgacggctgaagtcacagtcgccagaaacccacgtcatgccgttccgt 420
Qy 421 gcttgaagccgcgcgcgcagcatcccgagggggcatatccgaagcctcgtgcacgc 480
Db 421 gcttgaagccgcgcgcgcagcatcccgagggggcatatccgaagcctcgtgcacgc 480
Qy 481 gacgcctcgggtcgttggcgagcccgatgacagcgaccacgcctcttgaagccctgtgct 540
Db 481 gacgcctcgggtcgttggcgagcccgatgacagcgaccacgcctcgttgaagccctgtgct 540
Qy 541 ccaggacttcagcaggtgggtgtagagcgtggagcccaagtcctcgcgtggtggcg 600
Db 541 ccaggacttcagcaggtgggtgtagagcgtggagcccaagtcctcgcgtggtggcg 600
Qy 601 gggagacgtacacggtcgactcggcgtccagtcgtagcgttgcgtcctccagggc 660
Db 601 gggagacgtacacggtcgactcggcgtccagtcgtagcgttgcgtcctccagggc 660
```











Db	3540	attaaggtttcattg	tatttaattgtgtgcaaacatgg	cacttagtgtgtaggaaaaagtac	ca	3599
Qy	3601	aaatttgtctcaacctgatt	cagttcagttatggaattacattat	gaagctgtgctagagaa		3660
Db	3600	aaatttgtctcaacctgatt	cagttcagttatggaattacattat	gaagctgtgctagagaa		3659
Qy	3661	gatgtttattctagtcca	gcccacccacttatgcaagtc	tgtcttttagcttgattcaaaa		3720
Db	3660	gatgtttattctagtcca	gcccacccacttatgcaagtc	tgtcttttagcttgattcaaaa		3719
Qy	3721	actgatttaatttacat	tgtcctaattgtcacttc	cgagcctatgtcgcttttaattc	cgag	3780
Db	3720	actgatttaatttacat	tgtcctaattgtcacttc	cgagcctatgtcgcttttaattc	cgag	3779
Qy	3781	tagtagtgatatattag	tacataaaaaaatcatgtttg	aatcattcttctataaagtgc	acaa	3840
Db	3780	tagtagtgatatattag	tacataaaaaaatcatgtttg	aatcattcttctataaagtgc	acaa	3839
Qy	3841	gtcaattgtcccttctgt	tttggcactatattcaacttc	gtttaatgcaaaattatccag	tta	3900
Db	3840	gtcaattgtcccttctgt	tttggcactatattcaacttc	gtttaatgcaaaattatccag	tta	3899
Qy	3901	tacttagctagatatcca	attttgtaaaaaatagctct	gtatagtaaaacccgagtag	tg	3960
Db	3900	tacttagctagatatcca	attttgtaaaaaatagctct	gtatagtaaaacccgagtag	tg	3959
Qy	3961	acaaagtccatatccat	caaaacttcgtgtcgtggtc	gaagtcttgatcgacatggg	g	4020
Db	3960	acaaagtccatatccat	caaaacttcgtgtcgtggtc	gaagtcttgatcgacatggg	g	4019
Qy	4021	ttaaaatttaaat	tgggacacataaatagctct	tattgtgcataatccccatcg	aaaaatg	4080
Db	4020	ttaaaatttaaat	tgggacacataaatagctct	tattgtgcataatccccatcg	aaaaatg	4079
Qy	4081	acagattgtacatgaaa	aaaaaaagtcctctgata	gaagtgcgaagatcacaa	ttt	4140
Db	4080	acagattgtgtacatg	aaaaaaagtcctctgata	gaagtgcgaagatcacaa	ttt	4139
Qy	4141	tctatcgagagatagat	tgaaaagaagtcgaggaag	cgggttaactggaacataac	acacat	4200
Db	4140	tctatcgagagatagat	tgaaaagaagtcgaggaag	cgggttaactggaacataac	acacat	4199
Qy	4201	gtctaaatttaattgc	atctcgcttaacccaaaag	tgtattactctctcgggtcc	acataaag	4260
Db	4200	gtctaaatttaattgc	atctcgcttaacccaaaag	tgtattactctctcgggtcc	acataaag	4259
Qy	4261	ttatttttggccctttt	tattgttccaaaaataag	tgaagtttttttagatttca	aaaaaatg	4320
Db	4260	ttatttttggccctttt	tattgttccaaaaataag	tgaagtttttttagatttca	aaaaaatg	4319
Qy	4321	atttaatttttttact	acagtcgcttggagtaaat	ggtgtgtagtatgtgtaga		4380
Db	4320	atttaatttttttact	acagtcgcttggagtaaat	ggtgtgtagtatgtgtaga		4379
Qy	4381	aatgtttatgtgaagaa	tagtaaaaggttaattat	gatccaatttctatgtatt	taattgtt	4440
Db	4380	aatgtttatgtgaagaa	tagtaaaaggttaattat	gatccaatttctatgtatt	taattgtt	4439
Qy	4441	aaaaatgtaatttct	taattctgtgtaaaaaac	ccaaaaatcacttattgtg	gacccgga	4500
Db	4440	aaaaatgtaatttct	taattctgtgtaaaaaac	ccaaaaatcacttattgtg	gacccgga	4499
Qy	4501	gaaagtataaata	tatttttggagcgact	aaaaataaaacttttctcat	tattacga	4560
Db	4500	gaaagtataaata	tatttttggagcgact	aaaaataaaacttttctcat	tattacga	4559
Qy	4561	accttaaaaacag	catatgtgttctaggg	gaacttaactcaaaat	taataaagaa	4620
Db	4560	accttaaaaacag	catatgtgttctaggg	gaacttaactcaaaat	taataaagaa	4619
Qy	4621	gcaacaagtat	caatacatatgatt	tacacggtcaaacac	gaatttcgttaatt	4680
Db	4620	gcaacaagtat	caatacatatgatt	tacacggtcaaacac	gaatttcgttaatt	4679

Qy	4681	ataataagaattaatccaataagctcccaacccctataacttaactataaataaccagc	4740
Db	4680	ataataagaattaatccaataagctcccaacccctataacttaactataaataaccagc	4739
Qy	4741	gaatgatattatcatgacataatttatataataatgaatgataaactatgataatacaatgt	4800
Db	4740	gaatgatattatcatgacataatttatataataatgaatgataaactatgataatacaatgt	4799
Qy	4801	ataatctatgtatgttgtagaaaaagtagaaacaataataatagccgcgtatttgtataa	4860
Db	4800	ataatctatgtatgttgtagaaaaagtagaaacaataataatagccgcgtatttgtataa	4859
Qy	4861	aatccctaataataatgcgacgagatcccccggaattccgggaagcttagatccatgcag	4920
Db	4860	aatccctaataataatgcgacgagatcccccggaattccgggaagcttagatccatgcag	4919

RESULT 3

US-08-817-188-5

; Sequence 5, Application US/08817188

; Patent No. 6074876

; GENERAL INFORMATION:

; APPLICANT: DE BLOCK, MARC

; TITLE OF INVENTION: GENETIC TRANSFORMATION USING A PARP INHIBITOR

; FILE REFERENCE: 2121-0127P

; CURRENT APPLICATION NUMBER: US/08/817,188

; CURRENT FILING DATE: 1997-05-15

; EARLIER APPLICATION NUMBER: PCT/EP96/03366

; EARLIER FILING DATE: 1996-07-31

; EARLIER APPLICATION NUMBER: EP 95401844.6

; EARLIER FILING DATE: 1995-08-04

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5

; LENGTH: 5560

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: T-DNA of

; OTHER INFORMATION: plasmid pTHW142

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)..(25)

; OTHER INFORMATION: RB: right border sequence of octopine TL-DNA from

; OTHER INFORMATION: pTiB6S3

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (84)..(296)

; OTHER INFORMATION: 3' q7: 3' untranslated region containing the

; OTHER INFORMATION: polyadenylation signal of gene 7 of Agrobacterium

; OTHER INFORMATION: T-DNA

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (318)..(869)

; OTHER INFORMATION: bar: region coding for phosphinotricin

; OTHER INFORMATION: acetyltransferase

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (830)..(2760)

; OTHER INFORMATION: pSSU: promoter region of Rubisco small subunit

; OTHER INFORMATION: gene of Arabidopsis thaliana

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (2765)..(3058)

; OTHER INFORMATION: 3' untranslated region of the CamV 35S transcript

; OTHER INFORMATION: containing polyadenylation signals

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (3059)..(5056)

; OTHER INFORMATION: uidA: region coding for beta-glucuronidase

; FEATURE:

; NAME/KEY: misc\_feature

```

; LOCATION: (4483)..(4671)
; OTHER INFORMATION: IV2: region corresponding to the second intron of
; OTHER INFORMATION: the 5T-Ls1 gene
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5067)..(5502)
; OTHER INFORMATION: P35S: 35S promoter region of CaMV
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3533)..(5560)
; OTHER INFORMATION: LB: left border sequence of octopine TL-DNA from
; OTHER INFORMATION: PT1B6S3
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5058)..(5059)
; OTHER INFORMATION: region with unknown sequence (may contain up to 20
; OTHER INFORMATION: nucleotides)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5077)..(5078)
; OTHER INFORMATION: region with unknown sequence (may contain up to 20
; OTHER INFORMATION: nucleotides)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5476)..(5479)
; OTHER INFORMATION: region with unknown sequence (may contain up to 20
; OTHER INFORMATION: nucleotides)
; US-08-817-188-5

```

[illegible]

QY 1516 aatgtgtaagatagtcgcacatcaagtcacatctctgtttacaaatattgttacacaagaagtcata 1575  
Db 1667 aatgtgtaagatagtcgcacatcaagtcacatctctgtttacaaatattgttacacaagaagtcata 1726  
QY 1576 agcccaacaagttagcacgtctaaataaaactaaagaagtcacgaataattacaaatca 1635  
Db 1727 agcccaacaagttagcacgtctaaataaaactaaagaagtcacgaataattacaaatca 1786  
QY 1636 taagcccaacaagaattattgatcaaaaaaaaccccaaaagcctaaagcctaaacaaagtc 1695  
Db 1787 taagcccaacaagaattattgatcaaaaaaaaccccaaaagcctaaacaaagtc 1846  
QY 1696 aaaaaaaactctcaagtcctccatctctcttatgaacattgaaacatatcacaaaaa 1755  
Db 1847 aaaaaaaactctcaagtcctccatctctcttatgaacattgaaacatatcacaaaaa 1906  
QY 1756 agtcagataaactctctctggccgtctctccaaacctctacatcaactctccatcgg 1815  
Db 1907 agtcagataaactctctctggccgtctctccaaacctctacatcaactctccatcgg 1966  
QY 1816 attgaatgtttactgtacactttctccgttgcaatgatatgtatgtattgttgtaaaa 1875  
Db 1967 attgaatgtttactgtacactttctccgttgcaatgatatgtatgtattgttgtaaaa 2026  
QY 1876 ctaatagggttaacaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 1935  
Db 2027 ctaatagggttaacaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 2086  
QY 1936 ttctagtagaagaccatcacagaaatttactagtaaaataaaataaaataaaataaaat 1995  
Db 2087 ttctagtagaagaccatcacagaaatttactagtaaaataaaataaaataaaataaaat 2146  
QY 1996 ctattatgtgcgaataatcaataatataagaggaattttcaaatgaaaaacgctatgaat 2055  
Db 2147 ctattatgtgcgaataatcaataatataagaggaattttcaaatgaaaaacgctatgaat 2206  
QY 2056 gttattagtaaatggtcagtagaagacattaaaaaaatccctacgtcagatattcaacttta 2115  
Db 2207 gttattagtaaatggtcagtagaagacattaaaaaaatccctacgtcagatattcaacttta 2266  
QY 2116 aaaaatcgatcagtggaattgtacaaaatttggaatttggaatttggaatttggaatttgga 2175  
Db 2267 aaaaatcgatcagtggaattgtacaaaatttggaatttggaatttggaatttggaatttgga 2326  
QY 2176 ttcaaacacttggaatttttttggaagcgtggaatttttaactacatatattgttttggc 2235  
Db 2327 ttcaaacacttggaatttttttggaagcgtggaatttttaactacatatattgttttggc 2386  
QY 2236 catgcaccaactcattgttttagtgtaataactttgattttgtcgaataatattgttctgt 2295  
Db 2387 catgcaccaactcattgttttagtgtaataactttgattttgtcgaataatattgttctgt 2446  
QY 2296 atattgtataagaattctttgaccatatcacacacacatatatatatatatatatat 2355  
Db 2447 atattgtataagaattctttgaccatatcacacacacatatatatatatatatatat 2506  
QY 2356 attatatcatcgcacttttaattgaaaaaataatatatatatatatatattgtgcatttttc 2415  
Db 2507 attatatcatcgcacttttaattgaaaaaataatatatatatatattgtgcatttttc 2566  
QY 2416 taacaacacatatattgttcgcatgattcgaataaactcgtcgaagtaataagaaataata 2475  
Db 2567 taacaacacatatattgttcgcatgattcgaataaactcgtcgaagtaataagaaataata 2626  
QY 2476 atctattgctgaattatctcagatgttaagattttcttaagtaaaatttttcaaat 2535  
Db 2627 atctattgctgaattatctcagatgttaagattttcttaagtaaaatttttcaaat 2686  
QY 2536 tagctaaagctcttgaataactaaagaataataacaaactctcgaccacggaaaaaaac 2595  
Db 2687 tagctaaagctcttgaataactaaagaataataacaaactctcgaccacggaaaaaaac 2746  
QY 2596 acataataaatttgatt 2613

Db 2747 acataataaatttgatt 2764

RESULT 4  
US-08-104-072B-9  
; Sequence 9, Application US/08104072B  
; Patent No. 5639948  
; GENERAL INFORMATION:  
; APPLICANT: Michiels, Frank  
; APPLICANT: Morioka, Sinji  
; APPLICANT: Scheirlinck, Trees  
; APPLICANT: Komari, Toshiko  
; TITLE OF INVENTION: Stamen-specific Promoters from Rice  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merchant & Gould  
; STREET: 3100 No. 5639948west Center  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/104,072B  
; FILING DATE: 05-AUG-1993  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO 9200272  
; FILING DATE: 06-FEB-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 91403352.7  
; FILING DATE: 10-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 91402590.3  
; FILING DATE: 27-SEP-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 91400318.1  
; FILING DATE: 08-FEB-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kowalchuk, Katherine M.  
; REGISTRATION NUMBER: 36,848  
; REFERENCE/DOCKET NUMBER: 8076.93USWO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-332-5300  
; TELEFAX: 612-332-9081  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5620 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1..395  
; OTHER INFORMATION: /product= "pUC18 derived sequence"  
; FEATURE:  
; NAME/KEY: polyA\_site  
; LOCATION: 396..802  
; OTHER INFORMATION: /standard\_name= "from nopaline  
; OTHER INFORMATION: synthase gene from Agrobacterium T-DNA"  
; FEATURE:  
; NAME/KEY: promoter  
; LOCATION: 1139..1683  
; OTHER INFORMATION: /function= sequence derived from  
; OTHER INFORMATION: "tapetum-specific promoter of Nicotiana tabacum"  
; FEATURE:  
; NAME/KEY: CDS

```
; LOCATION: 803..1138
; OTHER INFORMATION: /product= "CDS of barnase gene"
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 1684..2515
; OTHER INFORMATION: /standard_name= "35S3 promoter"
; OTHER INFORMATION: sequence from Cauliflower mosaic virus Cabbb-J"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2517..3068
; OTHER INFORMATION: /product= "CDS of phosphinotricin
; OTHER INFORMATION: acetyltransferase gene"
; FEATURE:
; NAME/KEY: polyA_site
; LOCATION: 3069..3356
; OTHER INFORMATION: /standard_name= "Agrobacterium
; OTHER INFORMATION: T-DNA nopaline synthase gene"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3357..5620
; OTHER INFORMATION: /product= "pUC18-derived sequence"
; US-08-104-072B-9
```

```
Query Match 21.7%; Score 1270.4; DB 1; Length 5620;
Best Local Similarity 98.9%; Pred. No. 3e-268;
Matches 1291; Conservative 0; Mismatches 11; Indels 4; Gaps 1;
```

```
QY 2612 tttagccgcgcgtaccgcgggaattcgagctcggtaccgcgggagatctccgcatctagtaac 2671
Db TGTAAAGACGACGGCAGTAAATTCGAGCTCGGTACCGGGGATCTTCCCGATCTAGTAAC 437
QY 2672 atagatgacacgcgcgcgataattatcttagttgcgcgctatatatttggtttctatc 2731
Db ATAGATGACACCGCGCGGATAATTTATCTAGTTTTCGCGCTATATTTGTTTCTATC 497
QY 2732 gcgtattaatgtataattgcgggactctaatcataaaaccatctcataaatacgtc 2791
Db CGGTATTAATGTTAATTCGGGACCTAATCATATAAAACCCATCTCATATAATACGTC 557
QY 2792 atgattacattgtaattattacatgcttaacgttaatacagaataattatgataac 2851
Db ATGATTTACATGTTAATTTATATATACATGCTTAACGTAAATCAACAGAAATTTATGATAATC 617
QY 2852 atcgaaagacgcgcgaacagagttcaacttaagaacattttatgccaatgtttgaaga 2911
Db ATCGAAGACGCGCAACAGAGTTCAATCTTAAGAAACTTTATTGCCAAATGTTTGAACGA 677
QY 2912 tctgcttcgagtcctctagag---ccggaagtgaattgaccgcagatcagagttgaaga 2967
Db TCTGCTTCGGATCCTCTAGAGNNNCCGAAAGTGAATTTGACCGATCAGAGTTTGAAGA 737
QY 2968 aaaaattattacacactttatgtaagctgaagaaacgcgcgcgcgcgcgcgcgcgcgcgc 3027
Db AAAATTATTACACATTTATGTAAAGCTGAATAAAGCTGAAAGAAACGCGCTCCGACGAGCGGTTT 797
QY 3028 ttctgtatctgattttgttaaaggctctgataatggtcgcgtgtttgttaaatacagcag 3087
Db TTTCTGTTATCTGATTTTGTAAAGGCTGATATGCTGCTGCTGTTTGTAAATCAGCCAG 857
QY 3088 tcgcttgagtaagaatacgcggtctgaattctgaagcctgagttagttaataatccgcct 3147
Db TCCTTTGAGTAAGAATCCGGCTGAATTTCTCAAGCCCTGATGATAGTTAATATCCGCT 917
QY 3148 tcaagccattgttcgtccgctttttgcccggagtttgccctccctgtttgagagagtgct 3207
Db TCAGCCCATGTTTCGTCGCTTTTTCGCCGGAGTTTGCCTTCCTGTTTGAAGAGTGCT 977
QY 3208 ccgcgcgacttttccccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 3267
Db CGCCCATGCTTTTCCCGGAGCGAGCTGCAAGGTTCCCTTTTGTATGCCACCCAGCG 1037
QY 3268 agggctgtgctctcgattttgtaattgtaattcatcaggtagcttatgatgtctgaaga 3327
```

```
Db 1038 AGGCTTGTGCTCTGATTTTGTAAATGTAATATCAGTAGCTTATGATATGCTGAAGA 1097
QY 3328 taatccgcaaccccgctcaaacgctgttgataaccggtaccatgtagctaatcttttaag 3387
Db 1098 TAATCCGCAACCCCGTCAACGCTTGTGATACCGGTACCATGGTAGCTAATTTCTTTAAG 1157
QY 3388 taaaaacttgattgagtgatgattgtactgttactgtttacacttgcaccacagggcatata 3447
Db 1158 TAAAAACTTTGATTTGAGTGTGATGTTGTAC1GTTACACTTGCACCACAGGCGCATATA 1217
QY 3448 tagagcacaagacatacacaaacttgcaaaacttttggagcatttcgagga 3507
Db 1218 TAGAGCACAAAGACATACACAACTTGCAAAACTAAC1TTTGTGGAGCATTTTCGAGGA 1277
QY 3508 aatggggagtagcaggtctaatctgagggtaacatlaaggtttcatgtatttaattgttg 3567
Db 1278 AAATGGGAGTAGCAGGCTAATCTGAGGTAACATTAAGGTTTTCATGTTAATTTGTTG 1337
QY 3568 caaatatgacttagtgtagggaaaaagtaccaaaattttgtctaccctgatttcagtt 3627
Db 1338 CAAACATGGACTTAGTGTGAGGAAAAAGTACCAAAATTTTGTCTCACCTGATTTTCAGTT 1397
QY 3628 atggaattacattatgaagctgtgctadagaagatgtttattcttagtccagccaccac 3687
Db 1398 ATGGAATTTACATTTAAGAGCTGTGCTAGAGAAGATGTTTATTTCTAGTCCAGCCACCCAC 1457
QY 3688 ctatcagagctcgttttagcttgattcgaactgattcaaaactgatttaatttgcataaaatg 3747
Db 1458 CTTATGCAAGCTCTGCTTTTAGCTTGATTCAAAACTGA1TTAATTTACATTGCTAAATGT 1517
QY 3748 gcaacttcgagcctatgctgcctttaacttcgagtaggagtgatgtatatattagtcacataaa 3807
Db 1518 GCATATCTCGACCTATGTCGCTTTAATTCGAGTAGGATGATATATTAGTACATATAAAA 1577
QY 3808 atcatgttgaatcatctttcataaagtgcacagtcacaa1tgccctctctgtttggcact 3867
Db 1578 ATCATGTTTGAATCATCTTTTCATAAAGTCACAAGTCAAT1TGCCCTCTCTGTTTGGCACT 1637
QY 3868 atattcaactctgttaatgcaaatattccagttatcacttagctagat 3913
Db 1638 ATATTCATCTGTTAATGCAAAATTA1TCCAGTTATCTACTTAGCTAGAT 1683
```

```
RESULT 5
US-08-351-413-3
; Sequence 3, Application US/08351413
; Patent No. 5750867
; GENERAL INFORMATION:
; APPLICANT: Williams, Mark
; APPLICANT: Leemans, Jan
; TITLE OF INVENTION: Maintenance of male-sterile plants
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 2046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/351,413
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/899,072
; FILING DATE: 12-JUN-1992
; PRIOR APPLICATION DATA:
```

```

1 APPLICATION NUMBER: US 07/970,849
2 FILING DATE: 03-NOV-1992
3 ATTORNEY/AGENT INFORMATION:
4 NAME: Svensson, Leonard R.
5 REGISTRATION NUMBER: 30,330
6 REFERENCE/DOCKET NUMBER: 2121-102PCT
7 TELECOMMUNICATION INFORMATION:
8 TELEPHONE: (703) 205-8000
9 TELEFAX: (703) 205-8050
10 TELEX: 248345
11 INFORMATION FOR SEQ ID NO: 3:
12 SEQUENCE CHARACTERISTICS:
13 LENGTH: 5620 base pairs
14 TYPE: nucleic acid
15 STRANDEDNESS: double
16 TOPOLOGY: circular
17 MOLECULE TYPE: DNA (genomic)
18 HYPOTHETICAL: NO
19 ANTI-SENSE: NO
20 ORIGINAL SOURCE:
21 ORGANISM: plasmid pVE108 (replicable)
22 FEATURE:
23 NAME/KEY: -
24 LOCATION: 1..395
25 OTHER INFORMATION: /label= pUC18
26 OTHER INFORMATION: /note= "pUC18 der
27 FEATURE:
28 NAME/KEY: -
29 LOCATION: complement (396..802)
30 OTHER INFORMATION: /label= 3'nos
31 OTHER INFORMATION: /note= "3' regula
32 OTHER INFORMATION: polyadenylation s
33 OTHER INFORMATION: synthase gene fro
34 FEATURE:
35 NAME/KEY: -
36 LOCATION: complement (803..1138)
37 OTHER INFORMATION: /label= barnase
38 OTHER INFORMATION: /note= "coding re
39 OTHER INFORMATION: Bacillus amyloliq
40 FEATURE:
41 NAME/KEY: -
42 LOCATION: complement (1139..1683)
43 OTHER INFORMATION: /label= TA29
44 OTHER INFORMATION: /note= "sequence
45 OTHER INFORMATION: promoter of Nicot
46 FEATURE:
47 NAME/KEY: -
48 LOCATION: 1684..2516
49 OTHER INFORMATION: /label= 35S3
50 OTHER INFORMATION: /note= "35S3" pr
51 OTHER INFORMATION: cauliflower mosai
52 FEATURE:
53 NAME/KEY: -
54 LOCATION: 2517..3068
55 OTHER INFORMATION: /label= bar
56 OTHER INFORMATION: /note= "coding se
57 OTHER INFORMATION: acetyltransferase
58 FEATURE:
59 NAME/KEY: -
60 LOCATION: 3069..3356
61 OTHER INFORMATION: /label= 3'nos
62 OTHER INFORMATION: /note= "3' regula
63 OTHER INFORMATION: polyadenylation s
64 OTHER INFORMATION: T-DNA nopaline sy
65 FEATURE:
66 NAME/KEY: -
67 LOCATION: 3357..5620
68 OTHER INFORMATION: /label= pUC18
69 OTHER INFORMATION: /note= "pUC18 der
70 US-08-351-413-3

```

Best Local Similarity 98.98; Pred. No. 3e-268;				
Matches 1291; Conservative 0; Mismatches 11; Indels 4; Gaps 1;				
Qy	2612	tttcaccgcggtaccgcggaattcgcgctcgttaccgcggggtatctccgatctagtaac	2671	
Db	378	TGTAACAGCAGCGCCAGTGAATTCGAGCTCGTACCCGGGATCTTCCGATCTAGTAAC	437	
Qy	2672	atagatgacacccgcgcgataatttcctcagttttgcgcgtatatttggtttctatc	2731	
Db	438	ATAGATGACACCGCGCGGATAATTTATCTCTAGTTTGGCGCTATATTTTCTTCTATC	497	
Qy	2732	gcgtattaaatgtataattgcgggactcctaataaaaaaacccatctcataaaacgctc	2791	
Db	498	CGGTATTAATGTTAATATTCGGGACTTAATCATATAAAACCCATCTCATAAATAACGTC	557	
Qy	2792	atgcatcacatgttaattacatctcgtcttaacgttaattcaacagaaattatatataac	2851	
Db	558	ATGCATTACATGTTAATTAATTTACATGCTTAACGTAATTCACACAGAAATATATGATATC	617	
Qy	2852	atcgaagacgcgcaacaggattcacaattcctaagaacctttattgccaaatgtttgaacga	2911	
Db	618	ATCGCAAGACCGCAACAGGATTCAATCTTAAGAACTTTATTGCCAAATGTTTGAACGA	677	
Qy	2912	tctgcttcggaatcctctagag- - - ccggaagtgaaattgaccgatacagagtttgaaga	2967	
Db	678	TCTGCTTCGGATCCTCTAGAGNNNNCCGAAAGTGAAATTGACCGATCAGAGTTGAAGA	737	
Qy	2968	aaaatttattcacactttatgtaaacgtcaaaaaacggcctccgcagaagccgtttt	3027	
Db	738	AAAAATTAATTAACACTTTATGTAAAGCTGAATAATGGTCCGTGTTTGTAAATCAGCCAG	797	
Qy	3028	tttcgttactcgtattttgtaaaggtctgaataatggtcgtttgtttgtaaatcagccag	3087	
Db	798	TTTCGTTATCTGATTTTGTAAAGTCTGATAATGGTCCGTGTTTGTAAATCAGCCAG	857	
Qy	3088	tcgcttgagtaaaagatccgcgtcgtgaatttctgaagcctgatgatagtttaataatccgct	3147	
Db	858	TCGCTTGAGTAAAGAAATCCGCTCTGAATTTCTGAAGCCTGATGATAGTTAATATCCGCT	917	
Qy	3148	tcagccatgttcgcgcgttttcccgggagtttgcctccctcgttttgagaagatgctc	3207	
Db	918	TCAGCCCATGTTCGTGCGCTTTTCCCGGGAGTTTGCTTCCCTGTTTGAGAAAGATGCT	977	
Qy	3208	ccgcgcatgctttcccgggagcagcctgcaagggttcccttttgaagccaccagcgc	3267	
Db	978	CCGCGATGCTTTTCCCGGAGCGAGCTGCAAGGTCCCTTTTGATGCCACCCAGCGC	1037	
Qy	3268	agggcttggtcttgatatttgtaatgtaattatcaggtagcttatgatagtctgaaga	3327	
Db	1038	AGGGCTTGCTGTTCTGATTTTGTAAATGTAATATCAGGTAGCTTATGATATGCTGAAGA	1097	
Qy	3328	taatccgcaaacccgctcaaacgcttgataaccgggtaccatggtagcttaattctttaag	3387	
Db	1098	TAATCCGCAACCCCGTCAAAAGCTGTGATAACCGGTACCATGGTAGCTTAATTTCTTAAAG	1157	
Qy	3388	taaaactttgatgttgatgatgtttgactgttacaactgcaccacaagggcatata	3447	
Db	1158	TAAAAACTTTGATTTGATGATGATGTGTACTGTTACACTTGCACCACAGGCGATATA	1217	
Qy	3448	tagagcaacagacatacacacaacttcgaaacttttggtagagcatttcgagga	3507	
Db	1218	TAGAGCACAGACATACACAACACTTGCAAAACTAATTTTGTGGAGCATTTTCGAGGA	1277	
Qy	3508	aaatggggagtagcaggttaactcgtgagggtaaacattaaaggtttcatgtatttaattgttg	3567	
Db	1278	AAATGGGGAGTAGCAGGGTAACTCTGAGGGTAACATTAAGGTTTCATGTTAAATTTGTTG	1337	
Qy	3568	caacatggacttagtggaggaataacccaaatttctccacctgattcagtt	3627	
Db	1338	CAACATGGACTTAGTGTGAGGAATAAGTACCAAAATTTTGTCTCACCTCATTTTCAGTT	1397	
Qy	3628	atggaaattacattatgaagctgtgctagagaagatgtttatcttagtcagagccaccac	3687	

Query Match 21.7%; Score 1270.4; DB 1; Length 5620;

CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/351.413  
FILING DATE:  
APPLICATION NUMBER: US 07/899,072  
FILING DATE: 12-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/970,849  
FILING DATE: 03-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 2121-102PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5620 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: plasmid pVEI08 (replicable in E.coli)  
FEATURE:  
NAME/KEY: -

Query Match	21.7%	Score 1270.4	DB 2	Length 5620
Best Local Similarity	98.9%	Pred. No. 3e+368		
Matches 1291	Conservative 0	Mismatches 11	Indels 4	Gaps 1
QY 2612	tttcgacgcgcgtaccgcggaattcgagctcggtaccgcggggatcttcccgatcttagtaac	2671		
Db				
378	TGTAACAGCAGCGGCAGTGAATTTCGAGTCGGTACCGGGGATCTCCGATCTAGTAAC	437		
QY 2672	atagatgacacgcgcgcgataatttcctcagttttgcgcgtatatattgtttttctatc	2731		
Db				
438	ATAGATGACACCGCGCGGAGTAATTTACCTAGTTTGGCGCTATATTTGTTTCTATC	497		
QY 2732	gcgtattaaatgtataattgcgggactctaatcataaaacccatctcataaataacgc	2791		
Db				
498	CGCTATTAAATGTATAATTGGGGACTCTAATCATATAAAACCCATCTCATAAATAACGTC	557		
QY 2792	atgcattacatgttaattattacaagtcttaacgtaattcaacagagaattatgatgaatc	2851		
Db				
558	ATGCATTACATGTTTAATTATTACATGCTTACGTAATTCAACAGAAATTTATGATAATC	617		
QY 2852	atcgcgaagacgcgcacagagattcaactctaagaacctttatgccaaatgtttgaacga	2911		
Db				
618	ATCGCAAGACGGCAACAGGATTCATCTTAAAGAACTTTATGCCAAATGTTTGAACGA	677		



Qy 2912 tctgcttcgagatcctctagag-----ccggaaagtgaattgaccgatcagaggtttgaaga 2967  
Db 678 TCTCTCTCGGATCCTCAGAGNNNNCCGGAAGTGAATTGACCGATCAGAGTTTGAAGA 737  
Qy 2968 aaaatttatcacacatttatgaagctaaaaaacgacctccgaggaagccggtttt 3027  
Db 738 AAAATTTATTACACACTTTATGTAAGCTGAAAAAACGGCTCCGAGGAAGCGGTTTT 797  
Qy 3028 ttctgttatctgattttgttaaaggtctgataatgggtccgtgtgttttggtaaatcagccag 3087  
Db 798 TTTCTGTTATCTGATTTTGTGAAGGTCTGATAATGGTCCGTTGTTTGTAAATCAGCCAG 857  
Qy 3088 tcgcttgagtaaaagaaatccggctcgaatttctgaagcctgatgtatagtttaatatccgct 3147  
Db 858 TCGCTTGAGTAAGAAATCCGGCTCTGAATTTCTGAAGCCTGTATGTATATATCCGCT 917  
Qy 3148 tcagccatgtcttcgctcgtctttcccgaggagtttgcctccctgtttgagaagatgct 3207  
Db 918 TCAGGCCATGTTCTGCTCGCTTTTCCCGGGAGTTTGCCTTCCCTGTTTCAGAAATGTCT 977  
Qy 3208 ccgcccagatgttttcccgaggagcgtctgcaaggttcccttttgcaccaccagccg 3267  
Db 978 CCGCCGATGCTTTTCCCGGAGCGAGCTCTGCAAGGTTCCCTTTTGATGCCACCCAGCG 1037  
Qy 3268 agggcttgctctcgtatgtttgtaagttaataatcaggtagcttatgatatgctctgaaga 3327  
Db 1038 AGGGCTTGTGCTTCTGATTTTGTAAATGTAATATCAGGTAGCTTATGATATGCTGAAGA 1097  
Qy 3328 taatcccaaccccgtaaacgttttaataaccggtaccatggtagcttaattctttaag 3387  
Db 1098 TAATCCGCAACCCCGTCAACAGCTGTGATAACCGGTACCATGGTAGCTTAATTTCTTTAAG 1157  
Qy 3388 taaaaactttgatttgatgtatgtttgtactgttacacttgaccacaagggtcatata 3447  
Db 1158 TAAAAACTTTGATTTGAGTGATGTGTGTACTGTGTACACTTGCACCAAGGCGCATATA 1217  
Qy 3448 tagagcaagaacatacacacaacttgcaaaaactatttgttgagagcatttcgaaga 3507  
Db 1218 TAGAGCACAAAGACATACACAACAACCTTGCAAAAATAACTTTTGTGGAGCATTTTCGAGGA 1277  
Qy 3508 aaatggggagtagcaggtctgaaggttaacattgaaggtttcattgaattttgtt 3567  
Db 1278 AAATGGGGAGTAGCAGGCTAATCTGAGGGTAACATTAAAGTTTCATGTATTAAATTTGTTG 1337  
Qy 3568 caaacatggacttagtggagaaaaagtaacaaatttgcctcaccctgatttcagtt 3627  
Db 1338 CAACATGGACTTAGTGTGAGGAAAAAGTACCAAAATTTTCTCACCTGATTTCAGTT 1397  
Qy 3628 atggaattacattatgaagctgtctagagaagatgtttattctagtcagggcaccaccac 3687  
Db 1398 ATGGAATTTACATTATGAAGCTGTGCTAGAGAAGATGTTTATTCTAGTCCAGCCACCAC 1457  
Qy 3688 ctatgcaagctcgtcttttagcttgatttcaaaaactgaatttaattacattgctaaatgt 3747  
Db 1458 CTTATGCAAGTCTGCTTTTACGCTTGATTCAAAAACCTGAATTTAAATTTACATTGCTAAATGT 1517  
Qy 3748 gcatacttcgagcctatgtcgtctttaattcgagtaggagtataatttagtacataaaaa 3807  
Db 1518 GCATACTTCGAGCCTATGTCGCTTTAATTCAGTAGGATGTATATATTAGTACATAAAAA 1577  
Qy 3808 atcatgtttgaaatcatctttcataaagtacaaagtcgaattgcctcttctgttggcact 3867  
Db 1578 ATCATGTTTGAATCATCTTTTCATAAAGTGACAAGTCAATTTGCTCTTCTTGTGGCACT 1637  
Qy 3868 atattcaatcgttaagtcaaatatccagtttacttagctagat 3913  
Db 1638 ATATTCAATCTGTTAATGCAAAATTTATCCAGTTATACTTAGCTAGAT 1683

RESULT 7

US-08-064-121-3/c

; Sequence 3, Application US/08064121

; Patent No. 5641664

GENERAL INFORMATION:  
APPLICANT: D'HALLUIN, Kathleen  
APPLICANT: GOBEL, Elke  
TITLE OF INVENTION: PROCESS FOR TRANSFORMING MONOCOTYLEDONOUS PLANTS  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: George Mason Bldg., Washington & Prince Sts.  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/064,121  
FILING DATE: 24-MAY-1993  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 90403332.1  
FILING DATE: 23-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 91401888.2  
FILING DATE: 08-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Crane-Feury, Sharon E  
REGISTRATION NUMBER: 36,113  
REFERENCE/DOCKET NUMBER: 010830-043  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1287 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORGANISM: fragment of pTMT8  
FEATURE:  
NAME/KEY: -  
LOCATION: 1..545  
OTHER INFORMATION: /label= PTA29  
OTHER INFORMATION: /note= "Promoter from the TA29 gene of Nicotiana tabacum"  
FEATURE:  
NAME/KEY: -  
LOCATION: 546..881  
OTHER INFORMATION: /label= barnase  
OTHER INFORMATION: /note= "coding sequence of barnase gene"  
FEATURE:  
NAME/KEY: -  
LOCATION: 882..1287  
OTHER INFORMATION: /label= 3'nos  
OTHER INFORMATION: /note= "3' regulatory sequence containing the polyadenylation site derived from Agrobacterium T-DNA nopaline synthase gene"  
US-08-064-121-3

Query Match 21.6%; Score 1269; DB 1; Length 1287;  
Best Local Similarity 99.7%; Pred. No. 3.1e-268;  
Matches 1283; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

Qy 2631 aattcgagctcgggtaccgccgggggatcttcccgatctcagtaacatagatgacaccgcgcg 2690  
|||||

Db 1287 AATTGAGCTCGTACCGGGGATCTCCGATCTAGTAACATAGATGACACGCGCGG 1228  
 QY 2691 ataattactagtttgcgcgtatatttggtttctatcgcgtattataatgataatt 2750  
 Db 1227 ATAATTATCTAGTTGGCGGTATATTTGTTTCTATCCGTATTAATGTTAAAT 1168  
 QY 2751 ggggactataataaaccatctctataaatacgtcatgattacatgttaatta 2810  
 Db 1167 GCGGGACTTAATCAATAAAGCCATCTCATAAATAACGTATGATTTACATGTTAAITA 1108  
 QY 2811 ttacatgcttaacgtaattcaacagaaattatatataatcatcgcagagaccgcaacag 2870  
 Db 1107 TTACATGCTTAACGTAATTCACAGAAATATATGATATCATCATCAAGCGGCAACAG 1048  
 QY 2871 gattcaatttaagaacattttatgccaaatgtttgaacgatactgttcggaactctag 2930  
 Db 1047 GATTCAATCTTAAGAACTTTATGCCAAATGTTTGAACGATCTGCTTCGGATCTCTAG 988  
 QY 2931 ag----cggaagtgaaattgacgatacagagtttgagaataattttacacacttt 2986  
 Db 987 AGNNNCCGGAAGTGAATGACCGATCAGAGTTTGAAGAAATAATTTATACACACTTT 928  
 QY 2987 atgtaaagctgaaaaaacgagctccgcaggaagcgttttttgcgttatctgattttg 3046  
 Db 927 ATGTAAGCTGAAAAAACGGCTCCGAGGAGCGGTTTTTCGTTATCTGATTGTTG 868  
 QY 3047 taaaggtctgataaagtcggtgtgtttgttaaatcagccagtcgcttgataaagaatcc 3106  
 Db 867 TAAAGGTCTGATAATGTCGCTGTTGTTGTAATCAGCAGTCGCTTGAGTAAAGATCC 808  
 QY 3107 ggtctgaattttcgaagctgctgataatgataatcgcgttcacgcatgttcgctcgc 3166  
 Db 807 GGTCTGAATTTCTGAAGCCTGATGATAGTTAAATATCGCTTCACGCCATGTCGCTCGC 748  
 QY 3167 tttgcccgggagtttgctctccctgtttgagaagatgtctccgcgagtcgtttcccg 3226  
 Db 747 TTTTCCCGGGAGTTTGCTTCCCTGTTGAGAGATGCTCTCCGCCGATGCTTTCCCGC 688  
 QY 3227 gagcgagctcgaaggttccctttgatgccaccacgagcgaggtgtgtctctgatt 3286  
 Db 687 GAGGACGCTGCAAGGTTCCTTTGATGCCACCCAGCGGAGGCTTGCTCTGATT 628  
 QY 3287 ttgtaagttaattatcaggtagctttagatgtctgaagataatccgcaaccgcgcaaa 3346  
 Db 627 TTGTAATGTAATTTATCAGGTAGCTTATGATGCTGAAGATATCCGCAACCCGTCAA 568  
 QY 3347 acgtgttgataaacgggtaccatggtagctaatcttcttaagtaaaaaactttgattgagt 3406  
 Db 567 ACGTGTTGATAAACGGTACCATGGTAGCTAATTTCTTTAAAGTAAAAACTTTGATTGAGT 508  
 QY 3407 gatgatgttgactgttacacttgacacacagggcatatataagacacagacatacac 3466  
 Db 507 GATGATGTTGTTACTGTTACACTTGCACCAAGGGCATATATAGAGCAAGACATACAC 448  
 QY 3467 acaacttgcaaaactaacttttgbgagcattcagggaaatggggagtagcaggct 3526  
 Db 447 AACAACTTGCAAACTAACTTTTGTGGAGCAATTCGAGGAAATGGGAGTAGCAGGCT 398  
 QY 3527 aatcgaggtcaacattaaagtttcatgtattattttgttgcacacatggacttagtgy 3586  
 Db 387 AATCTGAGGTACATTAAAGTTTCATGTATTAATTTGTTGCAAAATGAGCTTAGTGTG 328  
 QY 3587 agaaaaaglaccaaaattttgtctcacctgatttcagttatggaattacattatgaa 3646  
 Db 327 AGGAAAAAGTACCAAAATTTGCTCACCTGATTTTCAGTTATGGAATACATTAAGAA 268  
 QY 3647 gctgtctagagaagatgtttatcttagtcagccaccacacctttagcaagtctctttt 3706  
 Db 267 GCTGTGCTAGAGAAGATGTTTATTTAGTCCAGGCCACCCACCTTATGCAAGTCTGCTTTT 208  
 QY 3707 agcttgattcaaaaactgatttaattacatttgctaaatgctacacttcgagcctatgt 3766  
 Db 207 AGCTTGATTCAAAACTGATTTAAATTTACATTGCTAAATGTCATACCTTCGAGCCTATGT 148

QY 3767 qccttaattogagtaggagtagtatattagcacataaaatcatgtttgaaatcattt 3826  
 Db 147 CGCTTTAATTCGAGTAGGATGATATATATTAGTACATAAAAAATCATGTTTGAATCATCTT 88  
 QY 3827 tcataaagtgacaagtcgaattgtccctctctgtttggcaactatattcaaatctgttaatgc 3886  
 Db 87 TCATAAAGTGACAAGTCAATTTGTCCCTTCTTGTGTCACCTATATTCAATCTGTTAATGC 28  
 QY 3887 aaattatccagtataacttagctagat 3913  
 Db 27 AATATATCCAGTTACTTAGCTAGAT 1

RESULT 8  
 US-08-478-015-3/c  
 ; Sequence 3, Application US/08478015  
 ; Patent No. 5712135  
 ; GENERAL INFORMATION:  
 ; APPLICANT: D'HALLUIN, Kathleen  
 ; APPLICANT: Gobel, Elke  
 ; TITLE OF INVENTION: PROCESS FOR TRANSFORMING  
 ; TITLE OF INVENTION: MONOCOTYLEDONOUS PLANTS  
 ; NUMBER OF SEQUENCES: 5  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Burns, Doane, Swecker & Mathis  
 ; STREET: P.O. Box 1404  
 ; CITY: Alexandria  
 ; STATE: Virginia  
 ; COUNTRY: United States  
 ; ZIP: 22313-1404  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: IBM PC compatible  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/478,015  
 ; FILING DATE: 07-JUN-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: EP 90403332.1  
 ; FILING DATE: 23-NOV-1990  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: EP 91401888.2  
 ; FILING DATE: 08-JUL-1991  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/064,121  
 ; FILING DATE: 23-JUN-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: McGowan, Malcolm K.  
 ; REGISTRATION NUMBER: 39,300  
 ; REFERENCE/DOCKET NUMBER: 010830-088  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703) 836-6620  
 ; TELEFAX: (703) 836-2021  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1287 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; HYPOTHEetical: NO  
 ; ANTI-SENSE: NO  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: fragment of pTMM8  
 ; FEATURE:  
 ; NAME/KEY: -  
 ; LOCATION: 1..545  
 ; OTHER INFORMATION: /label= pTA29  
 ; OTHER INFORMATION: /note= "Promoter from the TA29 gene of Nicotiana tabacum"  
 ; FEATURE:

; NAME/KEY: 546..881 /label= barnase
; LOCATION: 546..881
; OTHER INFORMATION: /note= "coding sequence of barnase gene"
; FEATURE:
; NAME/KEY: 882..1287
; LOCATION: 882..1287
; OTHER INFORMATION: /label= 3' nos
; OTHER INFORMATION: /note= "3' regulatory sequence containing the polyadenylation
; OTHER INFORMATION: site derived from Agrobacterium T-DNA nopaline synthase gene
; US-08-478-015-3

Query Match 21.6%; Score 1269; DB 1; Length 1287;
Best Local Similarity 99.7%; Pred. No. 3.1e-268;
Matches 1283; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

Qy 2631 aattcgagctcggtacccggggtatctcccgatctagtaacatagatgaacccgcgcgcg 2690
Db 1287 aattcgagctcggtacccggggtatctcccgatctagtaacatagatgaacccgcgcgcg 1228
Qy 2691 ataattatctatgctgctgctatatttcttctatcgctgattaaatgtataatt 2750
Db 1227 ATAATTTATCTAGTTGCGGCTATATTTTCTCTATCGCGTATTAATGTAATTT 1168
Qy 2751 gcgggactctaatcaataaaacccatctcataataaacgtcattacatgttaatta 2810
Db 1167 GCGGACTCTAATCAATAAAACCCATCTCATATAACGTCATGCAATTACATGTTAATTA 1108
Qy 2811 ttacatgcttaacgttaattcaacagaaattatatgataatcgaagccggcaacag 2870
Db 1107 TTACATGCTTAAACGTAATTCAACAGAAATTTATATGATAATCGCAAGACCGGCAACAG 1048
Qy 2871 gattcaatcttaagaaacttattgccaatgtttgacgatctgctcgatcctctag 2930
Db 1047 GATTCATCTTAAGAACTTTATGCGCAATGTTTGACCATGCTCTCGATCCTCTAG 988
Qy 2931 ag----ccggaaagtgaattgaccgcatcagagtttgaagaaaaattattacacacttt 2986
Db 987 AGNNNCCGGAAGTGAATGACCGATCAGAGTTTGAAGAAAAATTTATTACACACTTT 928
Qy 2987 atgtaagctgaaaaaaacccgctcccgaggaagccgttttctcgctatctgattttg 3046
Db 927 ATGTAAGAGCTGAAAAAACGGCTCCCGAGGAGCCGTTTTTTTCGTATCTGATTTTG 868
Qy 3047 taaaggtctgataatggtccgtttttttaaatacagccagtcgttgagtaaagaatcc 3106
Db 867 TAAAGGCTGATATGTCGCTGTTTTTGTAAATCAGCCAGTCGCTTGAGTAAAGATCC 808
Qy 3107 ggtctgaattctgaagcctgatgtatagtttaatatccgcttcacgcatgttcgctcgcg 3166
Db 807 GGTCTGAATTTCTGAAGCCTGATGATATAGTTAATATCCGCTTACGCGCATGTTCTGTCGC 748
Qy 3167 ttttcccgaggattgcttccctccctgtttgagaagatgtctcccgcatgctttcccg 3226
Db 747 TTTTCCCGGAGTTCCTTCCTGTTGAGAAAGATGTCTCCCGCATGCTTTTCCTCCCG 688
Qy 3227 gacgcgctctgaaggttcccttttgcacccagcagggctgtgcttctgatt 3286
Db 687 GAGCGAGCTGCAAGGTTCCTTTTGTATGCCACCCAGCCGAGGCTTGCTTCTGATTT 628
Qy 3287 ttgtaattgaattatcaggtagcttatgatgtctgaaagataatccgcaaccccgctcaa 3346
Db 627 TTGTAATGTAATATCAGGTAGCTTATGATATGCTGAAGATAATCCGCAACCCGCTCAA 568
Qy 3347 acgtgttgataacccggtaccatggagcctaatttcttaagtaaaaaacttttgattgagt 3406
Db 567 ACGTGTTGATAAACCGGTACCATTGCTAGCTAATTTCTTTAAGTAAAAAACTTTTGATTGAGT 508
Qy 3407 gatgatgttactgttacactgtcaccaagggcctatatagacacacacacacacac 3466
Db 507 GATGATGTTGACTGTTTACACTTTCACCAAGGGCATATATAGAGCAACAGACATACAC 448

Qy 3467 aacaactgcaaaactaacttttgttggagcatttcagagaaaaatggggagtagcaggct 3526
Db 447 AACAACTTGCAAAACTAACTTTTGTGGAGCATTTTCAGGAAAAATGGGAGTAGCAGGCT 388
Qy 3527 aatctgagggttaacattaaaggtttcatgtatttaatttggcaaacatggagcttagtgc 3586
Db 387 AATCTGAGGGTAACATTAAGGTTTCATGTATTAATTTGTCAAACATGGACTTTAGTGTG 328
Qy 3587 agaaaaagtcacaaaaattttgtctccacctgatttcagtttatgaaattacattatgaa 3646
Db 327 AGGAAAAAGTACCAAAATTTGTCTCACCCGTGATTTAGTTATGGAATTAACATTAAGAA 268
Qy 3647 gctgtgtagaagaagtgtttattctagtcagccacccaccttatgcaagtctgctttt 3706
Db 267 GCTGTGTAGAGAAAGATGTTTATTTCTAGTCCAGCCACCCACCTTATGCAAGTCTGCTTTT 208
Qy 3707 agcttgattcaaaactgatttaatttattacattgctaaatgtgcatactctgagcctatgt 3766
Db 207 AGCTTGATTCAAAAACATGATTTAATTTACATTTGCTAAATGTGCATCTTCGAGGCTATGT 148
Qy 3767 cgctttaattcagtagtagtgatataattagtcataaaaaaatcatgtttgaatcatctt 3826
Db 147 CGCTTTAATTCGAGTAGGATGATATATTAGTACATRAAAAAATCATGTTTGAATCATCTT 88
Qy 3827 tcataaagtgacaagtcgaattgtccctctctgtttggcaactatatcaaatctgtaatgc 3886
Db 87 TCATAAAGTGACAAGTCAATTTGCTCTTCTTTGTCACCTATATTCAATCTGTTAATGC 28
Qy 3887 aaattatccagttatacttagctagat 3913
Db 27 AAATATCCAGTTATACTTAGCTAGAT 1

RESULT 9
US-08-475-975-3/C
; Sequence 3, Application US/08475975
; GENERAL INFORMATION:
; APPLICANT: D'HALLUIN, Kathleen
; APPLICANT: GOBEL, Elke
; TITLE OF INVENTION: PROCESS FOR TRANSFORMING
; TITLE OF INVENTION: MONOCOTYLEDONOUS PLANTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: George Mason Bldg., Washington & Prince Sts.
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,975
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,121
; FILING DATE: 24-MAY-1993
; APPLICATION NUMBER: EP 90403332.1
; FILING DATE: 23-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91401888.2
; FILING DATE: 08-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 010830-043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620

```

; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1287 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: fragment of pTTH8
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..545 /label= PTA29
; OTHER INFORMATION: /note= "Promoter from the TA29 gene of Nicotiana
; OTHER INFORMATION: tabacum"
; FEATURE:
; NAME/KEY: -
; LOCATION: 546..881
; OTHER INFORMATION: /label= barnase
; OTHER INFORMATION: /note= "coding sequence of barnase gene"
; FEATURE:
; NAME/KEY: -
; LOCATION: 882..1287
; OTHER INFORMATION: /label= 3'nos
; OTHER INFORMATION: /note= "3' regulatory sequence containing the
; OTHER INFORMATION: polyadenylation site derived from Agrobacterium
; OTHER INFORMATION: T-DNA nopaline synthase gene "
US-08-475-975-3

Query_Match
Best Local Similarity 21.68; Score 1269; DB 3; Length 1287;
Matches 1283; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 2631 aattcgagctcgggacccgggagcttcccgatctagtaaacatagatgacacgcgcgcg 2690
Db 1287 AATTCGAGCTCGGTACCGGGGATCTCCGATCTAGTACATAGATGACACCGCGCGG 1228

QY 2691 ataattatcctagttgcgcgctatattttttctatcgcggtattaaatgataatt 2750
Db 1227 ATAATTATCTAGTTTGGCGGCTATATTGTTTCTATCGCGTATTAATGATATAAT 1168

QY 2751 gcgggacttaatacataaaacccatctataataacgcatgcatcatttaatt 2810
Db 1167 GCGGACTCTAATCATAAACCCATCTCATATAACGTCATGCAATACATGTTAATTA 1108

QY 2811 ttacatgcttaacgaattcaacagaaatttatatgataatcatcgcgaagaccgcaacag 2870
Db 1107 TTACATGCTTAAAGCTAATTCACAGAAATTTATATGATAATCATCGCAAGCGCAACAG 1048

QY 2871 gattcaatcttaagaactttattgcaaatggttgaagatgctggtcggatccctctag 2930
Db 1047 GATTCATCTTAGAAACTTTATTGGCAATGTTTGAAGCATCTGCTCGGATCCCTAG 988

QY 2931 ag----ccggaagtgaaattgaccgcatcagagtttgaaagaaaaatttatacaacttt 2986
Db 987 AGNNNNCCGGAAGTGAATTTGACCGATCAGAGTTTGAAGAAAAATTTATTACACTTT 928

QY 2987 atgtaagctgaaaaaagccctccgcgggaagccggttttttctgtatctgtatttg 3046
Db 927 ATGTAAGCTGAAAAAAGCGGCTCCGCGAGGAGCGGTTTTTTCGTTATCTGATTTTG 868

QY 3047 taaagctgataatggtcgttggtttgaatcagccagcagcgttgagtaaaatccc 3106
Db 867 TAAAGCTCTGATAATGTCGGTGTGTTTGAATCAGCCAGTCGCTTGAGTAAAGAAATCC 808

QY 3107 ggtcgaattctgaagcctgtagtatagttatcatccgcttcacgcatgcttcgcgcg 3166
Db 807 GGCTCGAATTTCTGAAGCCTGATGATAGTTAATATCCGCTTACGCCCATGTCGTCGCG 748

```

```

QY 3167 ttgtcccgaggagtttgccctccctgtttgagaagatgtotcccgcatgcttttccccc 3226
Db 747 TTTTCCCGGGAGTTTGCTTCCCTGTTGAGAGAGTCTCTCCCGCATGCTTTTCCCG 688

QY 3227 gagcagcgtctgcaaggttcccttttgcaccccccgcgagggcttgctctgatt 3286
Db 687 GAGCGACGTCGCAAGGTTCCCTTTTGCACCCACCGAGCGAGGGCTTGCTTCTGATT 628

QY 3287 ttgtaagttaattatccaggtagcttatgatgtctgaagataatccgcaacccgcgcaa 3346
Db 627 TTGTAATGTAATTATCAGGTAGCTTATGATATGCTCTGAAGATATCCGCAACCCGTCAA 568

QY 3347 acgtgttgataaacgggtaccatggtagctaatttctttaagtaaaactttgattgagt 3406
Db 567 ACGTGTGTAACCGGTACCATGGTAGCTAATTTCTTTAAGTAAAAAATTTGATTTGAGT 508

QY 3407 gatgatgttgactgttacacttgcaccacagggcatatatagacacacagacatacac 3466
Db 507 GATGATGTTGCTACTGTTACACTTGCACCACAGGGCATATATAGCACAAAGACATACAC 448

QY 3467 aacaacttgcaaaactaactttgttggagcatttcgaggaataatggggagtagcaggct 3526
Db 447 ARCAACTTGCAAAACCTAACTTTTGTGGAGCATTTTCAGGAAAAATGGGAGTAGCAGGCT 388

QY 3527 aatctgagggtaacattaaagtttccatgtatttaatttggcaaacatggacttagtgcg 3586
Db 387 AATCTGAGGGTACACTTAAGTTTCATGTATTAATTTGTGCAACATGGACTTAGTGTG 328

QY 3587 agaaaaagttaccaaaattttgtctccacctgatttcagttatggaaaattacattatgaa 3646
Db 327 AGGAAAAAGTACCAAAATTTTGTCTCACCTGATTTTCAGTTATGGAATTAATATATGAA 268

QY 3647 gctgtctagagaagatggtttatttattctagtcacagcccccacccattatgcagctctgt 3706
Db 267 GCTGTCTAGAGAAGATGTTTATTTCTAGTCCAGCCACCCACCTTATGCAAGTCTGCTTTT 208

QY 3707 agcttgattcaaaaactgatttaatttattcatgtctaaatgtgcatacttcgagcctatgt 3766
Db 207 AGCTTGATTCAAAACTGATTTAATTTACATTTGCTAAATGTGCATACTTCGAGCCCTATGT 148

QY 3767 cgttttaattcgcagtaggtgtatatatttagtcacataaaaaaatcatgtttgaatcatctt 3826
Db 147 CGCTTTAATTCGAGTAGGATGTATATATAGTACATAAAAAATCATGTTTGAATCATCTT 88

QY 3827 tcataaagtgaacaagccaattgcccctctctgttggcactatattcactatgtttaagc 3886
Db 87 TCATAAAGTGCAAGTCAATTTGCTCCCTTCTGTTTGGCACTATATTCATCTGTTAATGC 28

QY 3887 aaattatccagttatacttagctagtagat 3913
Db 27 AAATTATCCAGTTATCTTAGCTAGAT 1

```

```

RESULT 10
US-09-084-889-3/c
; Sequence 3, Application US/09084889
; Patent No. 6074877
; GENERAL INFORMATION:
; APPLICANT: D'HALLUIN, Kathleen
; APPLICANT: GOBEL, Elke
; TITLE OF INVENTION: PROCESS FOR TRANSFORMING
; TITLE OF INVENTION: MONOCOTYLEDONOUS PLANTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: George Mason Bldg., Washington & Prince Sts.
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/084,889

FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/064,121  
 FILING DATE:

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: EP 91401888.2  
 FILING DATE: 08-JUL-1991  
 ATTORNEY/AGENT INFORMATION:

NAME: Crane-Feury, Sharon E  
 REGISTRATION NUMBER: 36,113  
 REFERENCE/DOCKET NUMBER: 010830-043  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620  
 TELEFAX: (703) 836-2021  
 INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 1287 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double

TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO

ORIGINAL SOURCE:  
 ORGANISM: fragment of pTHW8  
 FEATURE:

NAME/KEY: -  
 LOCATION: 1..545  
 OTHER INFORMATION: /label= pTA29  
 OTHER INFORMATION: /note= "Promoter from the TA29 gene of Nicotiana tabacum"

FEATURE:  
 NAME/KEY: -  
 LOCATION: 546..881  
 OTHER INFORMATION: /label= barnase  
 OTHER INFORMATION: /note= "coding sequence of barnase gene"

FEATURE:  
 NAME/KEY: -  
 LOCATION: 882..1287  
 OTHER INFORMATION: /label= 3'nos  
 OTHER INFORMATION: /note= "3' regulatory sequence containing the polyadenylation site derived from Agrobacterium T-DNA nopaline synthase gene"

US-09-084-889-3

Query Match 21.6%; Score 1269; DB 3; Length 1287;  
 Best Local Similarity 99.7%; Pred. No. 3.1e-268;  
 Matches 1283; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

Qy 2631 aattcgagctcggtaccgaggatcttcgagatctagtaacatagatgacacccgcgcg 2690  
 |||  
 Db 1287 AATTCGAGCTCGGTACCGGGGATCTCCGATCTAGTAAATAGATGACACCCGCGCG 1228  
 |||  
 Qy 2691 ataattatcctagttcgcgctatattttgtttctatcgctattaaatgtataatt 2750  
 |||  
 Db 1227 ATAATTTATCTAGTTTCGCGCTATATTTTGTCTTCTATCGCGCTATTAAT 1168  
 |||  
 Qy 2751 qcgagacttaataataaaccatctcataaataacgltcatgacattgtaatta 2810  
 |||  
 Db 1167 GCGGAGCTTAATCATAAAACCCATCTCATATAACGTGATGATGATTAATTA 1108  
 |||  
 Qy 2811 ttacatgcttaacgtaattcaacagaattatataataatcgacacccgcaacag 2870  
 |||  
 Db 1107 TTACATGCTTAACGTAATCAACAGAAATATATGATAATCATCGCAAGCCGCAACAG 1048  
 |||  
 Qy 2871 gattcaatcttaagaactttattgccaatgtttgaaacgatctgctcgagctctag 2930  
 |||

Db 1047 GATTCAATCTTTAAGAAACCTTTTATTTGCCAAATGTTTGAACGATCTGCTTCGGATCCTCTAG 988  
 Qy 2931 ag----ccggaagtgaattgacccatcagagtttgaagaaaaattttattacacacttt 2986  
 |||  
 Db 987 AGNNNCCGGAAAGTGAATTTGACCGATCAGAGTTTGAAGAAAAATTTATTACACACTTT 928  
 |||  
 Qy 2987 atgtaaagctgaaaaaaacggcctcccgaggaagccggttttttcgttatctgattttg 3046  
 |||  
 Db 927 ATGTAAAGCTGAAAAAACGGCTCCGAGGAAGCCGTTTTTTTCGTTATCTGATCTTTTG 868  
 |||  
 Qy 3047 taaaggtctgataatggtccgtttttttttaaatacagccagctgctgagtaagaatcc 3106  
 |||  
 Db 867 TAAAGGCTGATAATGGTCCGTTGTTTGTAAATCAGCCATCGCTTGAGTAAAGAAATCC 808  
 |||  
 Qy 3107 ggtctgaatttctgaagcctgatgtatagtttaataatccgcttcacgcatgttcgtccgc 3166  
 |||  
 Db 807 GGTCTGAATTTCTGAAGCCTGATGATAGTTAATATCGGCTTCACGCCATGTTCTGTCGCG 748  
 |||  
 Qy 3167 tttgcccgggagtttgccttcctctgtttgagaagatgtctccgagatgcttctctctgatt 3226  
 |||  
 Db 747 TTTTCCCGGGAGTTTGCTTCCCTGTTTGAAGAGATGCTCCGCGCATGCTTTTCCCGCG 688  
 |||  
 Qy 3227 gacgagctctgcaaggttcccttttgaatgcccccagccgagggcttctctctctgatt 3286  
 |||  
 Db 687 GAGCGAGCTCTGCAAGGTTCCCTTTTGATGCCACCCAGCGGCTTTGTCTTCTTGAGT 628  
 |||  
 Qy 3287 ttgtaatgtaattatcaggttagcttatgatgtctgaagataatccgcaaccccgctcaa 3346  
 |||  
 Db 627 TTGTAATGTAATTTATCAGGTAGCTTATGATATGCTGAGATATATCCGCAACCCGCTCA 568  
 |||  
 Qy 3347 acgtgttgataacccggtaccatggtagctaaatttctttaaataaaactttgtattgagt 3406  
 |||  
 Db 567 ACGTGTGATTAACCGGTACCATGTTAGTAAATTTCTTTAAGTAAAAACTTTGATTGTAGT 508  
 |||  
 Qy 3407 gatgatgttgactgttacacttcaccacaaagggcatatatagagcaacagacataaac 3466  
 |||  
 Db 507 GATGATGTTGTACTGTTTACACTTGCACCAAGGGCATATATAGAGCAACAAGACATAGAC 448  
 |||  
 Qy 3467 acaacttgcaaaactaactttgttgagcatctcagagaaatggggagtagcagcgt 3526  
 |||  
 Db 447 AACAACTTGCAAAACTAACTTTTGTGAGCATTTTCAGAGAAATGGGAGTAGCAGGCT 388  
 |||  
 Qy 3527 aatcgagggtaacattaaaggtttcattgttaattttgttgcaaacatgagacttagtgtg 3586  
 |||  
 Db 387 AATCTGAGGGTAACATTAAGGTTTTCATGTAATTAATTTGTTGCAAACTAGGACTTAGTGTG 328  
 |||  
 Qy 3587 agaaaaagtaacaaaattttgtctcaccctgatttcagttatggaaattacattatgaa 3646  
 |||  
 Db 327 AGGAAAAAGTACCAAAATTTTGTCTCACCCCTGATTTTCAAGTTATGGAATTTACATTTATGAA 268  
 |||  
 Qy 3647 gctgtctagagaagatgtttattctagtcagccacccacttatgcaagctctgctttt 3706  
 |||  
 Db 267 GCTGTCTAGAGAGATGTTTATTTCTAGTCCAGCCACCCTTATGCAAGCTCTGCTTTT 208  
 |||  
 Qy 3707 agcttgattcaaaaactgatttaattacattgctaaatgctgacttcagagcctatgt 3766  
 |||  
 Db 207 AGCTTGATTCAAAACTGATTTAATTTACATTTGCTAAATGTGCATCTTCGAGCCTATGT 148  
 |||  
 Qy 3767 cgttttaattcogagtaggtatataatagtaacataaaataatcatgtttgaaatcatt 3826  
 |||  
 Db 147 CGCTTTAATTCGAGTAGGATGATATATTAGTACATAAAAAAATCATGTTTGAATCATCTT 88  
 |||  
 Qy 3827 tcataaagtgacaagtcgaattgctccttctgttggcaactatattcaaatctgtaagtc 3886  
 |||  
 Db 87 TCATAAAGTGACAGTCAATTTGTCCTTCTGTTGTTGGCACTATATTCAATCTGTTAATGC 28  
 |||  
 Qy 3887 aaattatccagttactacttagctagat 3913  
 |||  
 Db 27 AAATTATCCAGTTTACTTAGCTAGAT 1

```
; Sequence 2, Application US/08817188
; Patent No. 6074876
; GENERAL INFORMATION:
; APPLICANT: DE BLOCK, MARC
; TITLE OF INVENTION: GENETIC TRANSFORMATION USING A PARP INHIBITOR
; FILE REFERENCE: 2121-0127P
; CURRENT APPLICATION NUMBER: US/08/817,188
; EARLIER FILING DATE: 1997-05-15
; EARLIER APPLICATION NUMBER: PCT/EP96/03366
; EARLIER FILING DATE: 1996-07-31
; EARLIER APPLICATION NUMBER: EP 95401844.6
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 6548
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid pTS172
; NAME/KEY: misc_feature
; LOCATION: Complement((2019)..(2288))
; OTHER INFORMATION: 3' nos: 3' untranslated region containing the
; OTHER INFORMATION: polyadenylation signal of the nopaline synthase
; OTHER INFORMATION: gene of Agrobacterium T-DNA.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: Complement((2289)..(2624))
; OTHER INFORMATION: barnase: region coding for barnase
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: Complement((2625)..(4313))
; OTHER INFORMATION: PEI: promoter region of E1 gene of rice
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4336)..(5170)
; OTHER INFORMATION: P35S: 35S promoter region of Cauliflower Mosaic
; OTHER INFORMATION: Virus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: Complement((5711)..(6262))
; OTHER INFORMATION: bar: region coding for phosphinotricin
; OTHER INFORMATION: acetyltransferase
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (6263)..(6496)
; OTHER INFORMATION: 3'g': 3' untranslated region containing the
; OTHER INFORMATION: polyadenylation signal of gene 7 of Agrobacterium
; OTHER INFORMATION: T-DNA
; US-08-817-188-2

Query Match          14.2%; Score 832.8; DB 3; Length 6548;
Best Local Similarity 99.8%; Pred. NO. 1.4e-172;
Matches 834; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 54 gtacatggtgcataaagaaagcaattgttagatgttaattcccatcttgaagaataat 113
Db 6540 GTACATGGTCGATAGAAAAGGCAATTTGTAGATGTTAATTCCTCCATCTTGAAGAAATAT 6481

QY 114 agtttaaatatttattgataaaatacaagtcagggtattatgattcacaagcaaaaacataa 173
Db 6480 AGTTTAAATATTTATTGATAAAAATAACAAGTCAGGTATTATAGTCCAAAGCAAAACATAA 6421

QY 174 atttattgacgaagttaaattcagaataatttcaataaactgattatcatcagctgggtac 233
Db 6420 ATTTATTGATGCAAGTTTAAATTCAGAAATATTTCATAAATCACTGATTATATCAGCTGGTAC 6361

QY 234 attgcgtagatgaagactgagtcgcatattgtgttaatacataaaattgatgatag 293
Db 6360 ATTGCGGTAGATGAAGACTGAGTGGCATATTATGTGTAAATACATAAAATTGATGATATAG 6301
```

```
QY 294 ctagcttagctcatcgggggatcctagaacgcgtgactcagatctcgtgaagggcagg 353
Db 6300 CTAGCTTAGCTCATCGGGGATCTAGAACCGTGATCTCAGATCTCGGTGACGGGAGG 6241

QY 354 accggaagggcggtaccgcaggctgaagtccagctgccagaaacccacgctcatgcag 413
Db 6240 ACCGGAAGGGCGGTACCGGAGGCTGAAGTCCAGAGTCCAGAAACCCACGTCATGCCAG 6181

QY 414 ttcccgctgttgaagccgcccgcgcagcatccgcgcccggggggcattatccagagccctcg 473
Db 6180 TTCCCGTGTGTAAGCCGCGCCGCGAGCATGCCGGGGGGGCATATCCGAGCGCCTCG 6121

QY 474 tgaatgcacgcgtcgggtcgttgggcagcccgatgacagcagccacgctcttgaagccc 533
Db 6120 TGCATGCGCACGCTCGGGTGGTGGGCAGCCCGATGACGCGACCGCTCTTGAAGCCC 6061

QY 534 tgtctctccagggaacttcacagcgtgggttagagcgtggagccacgctcccgctcg 593
Db 6060 TGTCTCTCAGGGACTTCAGCAGGTGGGTGTAGAGCGTGGAGCCAGTCCCGTCCGCTGG 6001

QY 594 tggcgggggagacgtacacggtcgactcgccgctccagtcgtagcgttgctgccttc 653
Db 6000 TGGCGGGGGAGAGCTACACGGTCGACTCGGCCGTCCAGTCGTAGCGGTGCGTGCCTTC 5941

QY 654 caggggcccgctaggcgtccgcgcagcctccgctccacccctcgccgacgagccaggga 713
Db 5940 CAGGGCCCGCTAGCGCATGCCGGGACCTCGCCGTCCACCTCGCGCAGCAGCAGGGA 5881

QY 714 tagcgtcccgacagcagcagcagcgtcgctccactcctcggttccctcgcgctcgta 773
Db 5880 TAGCGCTCCCGCAGCAGCAGGAGTGGTTCGTCCTCCACTCCTCGGTTCTCGCGCTG 5821

QY 774 cggaagtgcacgctgttctctctctgtagtggtagcagatggtgcagaccccgccatg 833
Db 5820 CGGAAGTTGACCGTGTCTCGATGTAGTGTGTGACGATGCTGATGCTGAGCGTCCG 5761

QY 834 tcgcctcgtggcagcgcgagtcgcccggcgctcgcggcgctcgttctgggtccattgt 889
Db 5760 TCGCGCTCGTGGCAGCGGGGATGTCGGCGGGCGTCTGCTGTGGGTCCATGCTTAT 5705

RESULT 12
US-08-894-440-1/c
; Sequence 1, Application US/08894440
; Patent No. 6025546
; GENERAL INFORMATION:
; APPLICANT: PLANT GENETIC SYSTEMS N.V.
; TITLE OF INVENTION: Method to obtain male sterile plants
; FILE REFERENCE: NM50R
; CURRENT APPLICATION NUMBER: US/08/894,440
; CURRENT FILING DATE: 1997-11-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 6548
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid pTS174
; NAME/KEY: misc_feature
; LOCATION: (1)..(2003)
; OTHER INFORMATION: pUC19 derived vector sequences (vector)
; NAME/KEY: misc_feature
; LOCATION: Complement((2019)..(2283))
; OTHER INFORMATION: 3' nos: region containing polyadenylation signal
; OTHER INFORMATION: of nopaline synthase gene of Agrobacterium
; NAME/KEY: misc_feature
; LOCATION: Complement((2284)..(2624))
; OTHER INFORMATION: region coding for barnase of Bacillus
; OTHER INFORMATION: amyloliquefaciens
```

```

;
; NAME/KEY: misc_feature
; LOCATION: Complement((2625)..(4313))
; OTHER INFORMATION: promoter of the stamen-specific E1 gene of rice
; OTHER INFORMATION: (PE1)
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4336)..(5710)
; OTHER INFORMATION: 35S promoter of Cauliflower Mosaic Virus (P35S)
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5711)..(6262)
; OTHER INFORMATION: region coding for phosphinothricin acetyl
; OTHER INFORMATION: transferase (bar)
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (6263)..(6496)
; OTHER INFORMATION: region containing polyadenylation signal fo gene 7
; OTHER INFORMATION: of Agrobacterium T-DNA (3'g7)
;
US-08-894-440-1

Query Match          14.2%; Score 831.2; DB 3; Length 6548;
Best Local Similarity 99.6%; Pred. No. 3.1e-172;
Matches 833; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 54 gtacatggtcgataaagaaaggcaattgtagatgtttaattcccatcttgaagaaatat 113
Db 6540 GTACATGTCGATAAGAAAGGCAATTGTAGATGTTAAATCCCATCTTGAAGAAATAT 6481

QY 414 agtttaaatattattgataaaataacaagtcagggtattatagtcacgaagcaaaacataa 173
Db 6480 AGTTTAAATATTTATTGATAAAATAACAAGTCAGGTATTATAGTCCAAAGCAAAACATAA 6421

QY 174 atttattgatcgaagttaaattcaaaataatttcaataactaattatatacagctgctac 233
Db 6420 ATTTATTGATGCAAGTTTAATTAATTCAGAAATATTTCATAACTGATTATATACAGCTGTGATC 6361

QY 234 attgcgtagatgaagagactgagtcgcatattatgttaatacatataaattgatgatag 293
Db 6360 ATTGCCGTAGATGAAGACTGAGTGCATATTATGTGTAATACATAAAATTTGATGATATAG 6301

QY 294 ctagcttagctatcgggggatcctagaacgctgtagctcagatctcgttgacggggcagg 353
Db 6300 CTAGCTTAGCTATCGGGGGATCCTAGAACGCTGTCTCAGATCTCGGTGACGGGCAGG 6241

QY 354 accgagcgggctgacccgagctgaagtcagctgagtcgagcaaaaccacgtcatgccag 413
Db 6240 ACCGGACGGGGCGGTACCGCGAGCGCTGAAGTCCAGCTGCCAGAAACCCACGTCATGCCAG 6181

QY 414 ttccctgtctgaagccgcccgcgcagcatgccgagggggcatatccgagcgcctcg 473
Db 6180 TTCCCGTGTGTTGAAGCGGGCCGCCGAGCATGCCGGGGGGCATATCCGAGCGCCTCG 6121

QY 474 tgatgcgcacgctcgggtcgcttgggcagcccgatgacagcagccacgctcttgaagccc 533
Db 6120 TGATGCGCACGCTCGGGTCTGTTGGGAGCCCGCATGACAGCACGACGCTCTTTGAAGCCC 6061

QY 534 tgtgcctccagggacttcagcaggtgggtgtagagcgtgagccagtcgccgtccgctcg 593
Db 6060 TGTGCCCTCCAGGAGACTTCACAGGTGGGTGTAGAGCGTGAGGCCAGTCCCGTCCGCTTG 6001

QY 594 tggcgggggagacgtacacaggtcgactcggccgctccagtcgtagggcgttgctgcttc 653
Db 6000 TGCGGGGGGAGAGCGTACACGGTTCGACTCGCGCGTCCAGTCGTAGGCGTTGCGTGCCTTC 5941

QY 654 caggggcccgctagcgatgacgagcagcctcgcctccacacctggcgacagcagcgga 713
Db 5940 CAGGGGCCCGGTAGGCGATGCCGGCACCTCGCCGCTCCACCTCGCGGACGAGCGAGGGA 5881

QY 714 taacgctcccgacgagcagagtgcttcgctccactccctgcggttcctcggtcggtgta 773
Db 5880 TAGCGCTCCCGACAGAGGAGGTGCTGCTCCGCTCCACTCTTCCGCTGCGGCTCGGTA 5821
```

```

QY 774 cggaaattgaccgtgcttctcgtatgtagtggtagcagtggtgacgacccgagccgcatg 833
Db 5820 CGGAAGTTGACCGTGTGCTCGATCTAGTGGTTGACGATGGTGACACCGCGCGCATG 5761

QY 834 tcgagctcggtagcagcggcgagatgctggccggcgctcgttctcgggtccattgttct 889
Db 5760 TCCGCCCTGGGTGGCAGCGCGGATGTCGGCGGGCGTCTGTTCTGGGTCCATGGTTAT 5705

RESULT 13
US-08-894-440-3
; Sequence 3, Application US/08894440
; Patent No. 6025546
; GENERAL INFORMATION:
; APPLICANT: PLANT GENETIC SYSTEMS N.V.
; TITLE OF INVENTION: Method to obtain male sterile plants
; FILE REFERENCE: WMSOR
; CURRENT APPLICATION NUMBER: US/08/894,440
; CURRENT FILING DATE: 1997-11-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 3658
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: EcoRI-HindIII
; OTHER INFORMATION: fragment of pVE136
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(26)
; OTHER INFORMATION: polylinker of pUC19
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: Complement((28)..(403))
; OTHER INFORMATION: region containing polyadenylation signal of
; OTHER INFORMATION: napaline synthase gene of Agrobacterium T-DNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: Complement((404)..(739))
; OTHER INFORMATION: region coding for barnase of Bacillus
; OTHER INFORMATION: amyloliquefaciens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: Complement((740)..(1918))
; OTHER INFORMATION: promoter of CA55 gene of Zea mays (PCA55)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1956)..(2788)
; OTHER INFORMATION: 35S promoter of Cauliflower Mosaic Virus (35S)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2789)..(3340)
; OTHER INFORMATION: region coding for phosphinothricin acetyl
; OTHER INFORMATION: transferase (bar)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3341)..(3623)
; OTHER INFORMATION: region containing polyadenylation signal of
; OTHER INFORMATION: napaline synthase gene of Agrobacterium T-DNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3624)..(3658)
; OTHER INFORMATION: polylinker of pUC19
;
US-08-894-440-3
```

Query Match 12.6%; Score 741.8; DB 3; Length 3658;  
Best Local Similarity 99.7%; Pred. No. 8.2e-153;  
Matches 743; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 2630 gaattcagctggtaccggggtatcttcctgcatctagtaacatagatgacacccgcgc 2689
D 1 gaattcagctggtaccggggtatcttcctgcatctagtaacatagatgacacccgcgc 60
QY 2690 gataattatctctagtttcgcgcgtatattttgtttctatcgctgattaaatgtataat 2749
D 61 gataattatctctagtttcgcgcgtatattttgtttctatcgctgattaaatgtataat 120
QY 2750 tgcgggactctaatcataaaacccatctcataaaacgcgcgcgcgcgcgcgcgcgcgcgc 2809
D 121 tgcgggactctaatcataaaacccatctcataaaacgcgcgcgcgcgcgcgcgcgcgcgc 180
QY 2810 attaatgcttaacgtaattcaacagaaattatgataatcatcgcgaacccgcgcgcgcgc 2869
D 181 attaatgcttaacgtaattcaacagaaattatgataatcatcgcgaacccgcgcgcgcgc 240
QY 2870 gattcaactcttaagaacactttattgcaaatgtttgaaacgatctgctcgatctctcta 2929
D 241 gattcaactcttaagaacactttattgcaaatgtttgaaacgatctgctcgatctctcta 300
QY 2930 gagcgggaagtgaattgaacgacagagtttgaagaaattattacacactttatg 2989
D 301 gagcgggaagtgaattgaacgacagagtttgaagaaattattacacactttatg 360
QY 2990 taaagctgaacaaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 3049
D 361 taaagctgaacaaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 420
QY 3050 aggtctgataatggtccgttttttgaataacccagcagcgcgcgcgcgcgcgcgcgcgcgc 3109
D 421 aggtctgataatggtccgttttttgaataacccagcagcgcgcgcgcgcgcgcgcgcgcgc 480
QY 3110 ctgaattctgaacgctgagtgatagtaattatccgcttcacgcgcgcgcgcgcgcgcgcgcgc 3169
D 481 ctgaattctgaacgctgagtgatagtaattatccgcttcacgcgcgcgcgcgcgcgcgcgcgc 540
QY 3170 tgcgcggagtttgcctccctcttgaagagatgctccgcgcgcgcgcgcgcgcgcgcgcgcgc 3229
D 541 tgcgcggagtttgcctccctcttgaagagatgctccgcgcgcgcgcgcgcgcgcgcgcgcgc 600
QY 3230 cgaactgaaaggtccctttgatgccacccagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 3289
D 601 cgaactgaaaggtccctttgatgccacccagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 660
QY 3290 taatgaattatcaggtagcttatgatctgtaagataatccgcaacccgcgcgcgcgcgcgcgc 3349
D 661 taatgaattatcaggtagcttatgatctgtaagataatccgcaacccgcgcgcgcgcgcgcgc 720
QY 3350 tgttgataacccggtaccatggtgacg 3374
D 721 tgttgataacccggtaccatggtgacg 745
```

## RESULT 14

US-08-351-413-17

; Sequence 17, Application US/08351413

; Patent No. 5750867

; GENERAL INFORMATION:

; APPLICANT: Williams, Mark

; APPLICANT: Leemans, Jan

; TITLE OF INVENTION: Maintenance of male-sterile plants

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESSES:

; ADDRESSES: BIRCH, STEWART, KOLASCH &amp; BIRCH

; STREET: 8110 Gatehouse Road, Suite 500 East

; CITY: Falls Church

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 2046

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/351,413
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/899,072
; FILING DATE: 12-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/970,849
; FILING DATE: 03-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Svensson, Leonard R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 2121-102PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4808 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: EcoRI-HindIII fragment of plasmid pTS218
; FEATURE:
; NAME/KEY: -
; LOCATION: complement (18..401)
; OTHER INFORMATION: /label= 3'nos
; OTHER INFORMATION: /note= "3' regulatory sequence containing the
; OTHER INFORMATION: polyadenylation site derived from Agrobacterium
; OTHER INFORMATION: T-DNA nopaline synthase gene"
; FEATURE:
; NAME/KEY: -
; LOCATION: complement (402..737)
; OTHER INFORMATION: /label= barnase
; OTHER INFORMATION: /note= "coding region of the barnase gene of
; OTHER INFORMATION: Bacillus amyloliquefaciens"
; FEATURE:
; NAME/KEY: -
; LOCATION: complement (738..1944)
; OTHER INFORMATION: /label= PZM13
; OTHER INFORMATION: /note= "promoter region of the Zm13 gene of Zea
; OTHER INFORMATION: mays"
; FEATURE:
; NAME/KEY: -
; LOCATION: complement (1945..2281)
; OTHER INFORMATION: /label= 3'nos
; FEATURE:
; NAME/KEY: -
; LOCATION: complement (2282..2554)
; OTHER INFORMATION: /label= barstar
; OTHER INFORMATION: /note= "coding region of the barstar gene of
; OTHER INFORMATION: Bacillus amyloliquefaciens"
; FEATURE:
; NAME/KEY: -
; LOCATION: complement (2555..3099)
; OTHER INFORMATION: /label= PTA29
; OTHER INFORMATION: /note= "promoter region of the TA29 gene of
; OTHER INFORMATION: Nicotiana tabacum"
; FEATURE:
; NAME/KEY: -
; LOCATION: 3100...3932
; OTHER INFORMATION: /label= 35S3
; OTHER INFORMATION: /note= "35S3* promoter sequence derived from
; OTHER INFORMATION: cauliflower mosaic virus isolate CabBB-JI"
; FEATURE:
```



```

: NAME/KEY: -
: LOCATION: 3933...4484
: OTHER INFORMATION: /label= bar
: OTHER INFORMATION: /note= "coding region of the phosphinothricin
: OTHER INFORMATION: acetyltransferase gene"
: FEATURE:
: NAME/KEY: -
: LOCATION: 4485...4763
: OTHER INFORMATION: /label= 3'nos
: FEATURE:
: NAME/KEY: -
: LOCATION: 2333...2356
: OTHER INFORMATION: /label= BXOL2
: OTHER INFORMATION: /note= "region corresponding to oligonucleotide
: OTHER INFORMATION: BXOL2"
: FEATURE:
: NAME/KEY: -
: LOCATION: complement (2538..2586)
: OTHER INFORMATION: /label= TA29SBXOL2
: OTHER INFORMATION: /note= "region complementary to oligonucleotide
: OTHER INFORMATION: TA29SBXOL2"
: FEATURE:
: NAME/KEY: -
: LOCATION: complement (2800..2823)
: OTHER INFORMATION: /label= PTA29OL5
: OTHER INFORMATION: /note= "region complementary to part of
: OTHER INFORMATION: oligonucleotide PTA29OL5"
: US-08-351-413-17

Query Match          12.4%; Score 729.4; DB 1; Length 4808;
Best Local Similarity 99.3%; Pred. No. 4.8e-150;
Matches 744; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

Qy 2630 gaattcagctcggtaccggggtctcccgatctagtaacatagatgacaccgcgcg 2689
Db 1 GAATTCGAGCTCGGTACCCGGGGATCTCCCGATCTAGTACATAGATGACACCGCGCGC 60

Qy 2690 gataattatctagttggcgcgctatatattgttttctatcgcgattaaatgataat 2749
Db 61 GATAATTATCTAGTTGGCGCGCTATATTGTGTTCTATCGGTATTAATGTATAAT 120

Qy 2750 tgcgggactctaatcaaaaaacccatctcataataaacgtcatgcatatttaatt 2809
Db 121 TCGGGGACTCTAAATCAATAAAACCCATCTCATATAATACGTCATGCATTTACATTTAAT 180

Qy 2810 attacatgcttaacgtaattcaacagaaattatgataatcatcgaaaccggcaaca 2869
Db 181 ATTACATGCTTAACGTAATTCAACAGAAATTTATGATAATCATCGCAAGACCGCAACA 240

Qy 2870 ggattcaatctgaagaactttattgcaaatgtttgaacgatctgctcgatctctcta 2929
Db 241 GGATTCAATCTTAAGAAACTTTATTGCAAAATGTTTGAACGATCTGCTTCGGATCCTCTA 300

Qy 2930 gag----ccggaaagtgaattgaccgatcagagtttgaagaaaaattattacacatt 2985
Db 301 GAGNNNCCGGAAAGTGAATTTGACCATCAGAGTTTGAAGAAAAATTTATACACACTT 360

Qy 2986 tatgaagctgaaaaaacggctccgcaggaagcggttttttctgattctgattttt 3045
Db 361 TATGTAAGCTGAAAAAANACGGCCCTCCGAGGAAGCCGTTTTTTTCGTATCTGATTTTT 420

Qy 3046 gtaaggtctgataatgggtccgtttgtttgtaatacagcagtcgcttgagtaagaatc 3105
Db 421 GTAAAGCTCTGATAATGGTCGTTGTTTGTGTAATACGCCAGTCGCTTGAGTAAAGAAATC 480

Qy 3106 cggctgaattctgaagcctgatgtatagtttaataatccgttcacgccattgttcgcg 3165
Db 481 CGGTCTGAATTTCTGAAGCCTGATGTATAGTTAATAATCCGCTTCACGCCATGTTTCGTCG 540

Qy 3166 cttttgccgggagtttgcttcccttttgagagaatgctccgcgagatgcttttccc 3225
Db 541 CTTTGGCCGGGAGTTTGGCCTTCCCTGTTTGAGAAGATGCTCCGCCGATGCTTTTCCCC 600
```

```

Qy 3226 ggagcgacgtctgcaaggtttcccttttgatgccaccagcgaggggttgcttctgat 3285
Db 601 GGAGCGACGCTGCGAAGGTTCCCTTTTGTATGCCACCGAGCGAGGCTTGTGCTTCTGAT 660

Qy 3286 ttgttaatgaattatcaggttagcttatgatattctgaagataatccgcaaccccgatca 3345
Db 561 TTTGTAATGTAATATATCAGGTAGCTTATGATATGCTGAAGATATCCGCAACCCCGTCA 720

Qy 3346 aacgtgttgataaccggtaccatggtagc 3374
Db 721 AACGTGTTGATAACCGGTACCATGGTTGC 749

RESULT 15
US-09-583-17
: Sequence 17, Application US/09025583
: Patent No. 5977433
: GENERAL INFORMATION:
: APPLICANT: Williams, Mark
: APPLICANT: Leemans, Jan
: TITLE OF INVENTION: Maintenance of male-sterile plants
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
: STREET: 8110 Gatehouse Road, Suite 500 East
: CITY: Falls Church
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 2046
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/025,583
: FILING DATE:
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/351,413
: FILING DATE:
: APPLICATION NUMBER: US 07/899,072
: FILING DATE: 12-JUN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/970,849
: FILING DATE: 03-NOV-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Svensson, Leonard R.
: REGISTRATION NUMBER: 30,330
: REFERENCE/DOCKET NUMBER: 2121-102PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 205-8000
: TELEFAX: (703) 205-8050
: TELEX: 248345
: INFORMATION FOR SEQ ID NO: 17:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4808 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: EcoRI-HindIII fragment of plasmid pTS218
: FEATURE:
: NAME/KEY: -
: LOCATION: complement (18..401)
: OTHER INFORMATION: /label= 3'nos
: OTHER INFORMATION: /note= "3' regulatory sequence containing the
: OTHER INFORMATION: polyadenylation site derived from Agrobacterium
: OTHER INFORMATION: T-DNA nopaline synthase gene"
```

```

FEATURE:
  NAME/KEY: -
  LOCATION: complement (402..737)
  OTHER INFORMATION: /label= barnase
  OTHER INFORMATION: /note= "coding region of the barnase gene of
  OTHER INFORMATION: Bacillus amyloliquefaciens"
FEATURE:
  NAME/KEY: -
  LOCATION: complement (738..1944)
  OTHER INFORMATION: /label= P2M13
  OTHER INFORMATION: /note= "promoter region of the Zm13 gene of Zea
  OTHER INFORMATION: mays"
FEATURE:
  NAME/KEY: -
  LOCATION: complement (1945..2281)
  OTHER INFORMATION: /label= 3'nos
FEATURE:
  NAME/KEY: -
  LOCATION: complement (2282..2554)
  OTHER INFORMATION: /label= barstar
  OTHER INFORMATION: /note= "coding region of the barstar gene of
  OTHER INFORMATION: Bacillus amyloliquefaciens"
FEATURE:
  NAME/KEY: -
  LOCATION: complement (2555..3099)
  OTHER INFORMATION: /label= PTA29
  OTHER INFORMATION: /note= "promoter region of the TA29 gene of
  OTHER INFORMATION: Nicotiana tabacum"
FEATURE:
  NAME/KEY: -
  LOCATION: 3100..3932
  OTHER INFORMATION: /label= 35S3
  OTHER INFORMATION: /note= "35S3" promoter sequence derived from
  OTHER INFORMATION: cauliflower mosaic virus isolate CabbB-JI"
FEATURE:
  NAME/KEY: -
  LOCATION: 3933..4484
  OTHER INFORMATION: /label= bar
  OTHER INFORMATION: /note= "coding region of the phosphinothricin
  OTHER INFORMATION: acetyltransferase gene"
FEATURE:
  NAME/KEY: -
  LOCATION: 4485..4763
  OTHER INFORMATION: /label= 3'nos
FEATURE:
  NAME/KEY: -
  LOCATION: 2333..2356
  OTHER INFORMATION: /label= BXOL2
  OTHER INFORMATION: /note= "region corresponding to oligonucleotide
  OTHER INFORMATION: BXOL2"
FEATURE:
  NAME/KEY: -
  LOCATION: complement (2538..2586)
  OTHER INFORMATION: /label= TA29SBXOL2
  OTHER INFORMATION: /note= "region complementary to oligonucleotide
  OTHER INFORMATION: TA29SBXOL2"
FEATURE:
  NAME/KEY: -
  LOCATION: complement (2800..2823)
  OTHER INFORMATION: /label= PTA29OL5
  OTHER INFORMATION: /note= "region complementary to part of
  OTHER INFORMATION: oligonucleotide PTA29OL5"

```

US-09-025-583-17

Query Match 12.4%; Score 729.4; DB 2; Length 4808;  
 Best Local Similarity 99.3%; Pred. No. 4.8e-150;  
 Matches 744; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

```

QY 2630 gaattcgagtcggtaccgggaggtctcccgctctagtaacatagatgacacccgcgc 2689
  |||||
Db 1 GAATTCGAGTCGTCGACCCGGGAGTCTCCCGATCTAGTAACATAGATGACACCCGCGC 60

```

```

QY 2690 gataattatcctagtttgcgcgcgtatatttggttttctatcgcgtattaaatgataat 2749
  |||||
Db 61 GATAAATTATCTAGTTTGCgcgcgtATATTGTTTCTATCGGTAATTAAATGATAAAT 120
QY 2750 tgcgggactctaataaaacccatctcataaataacgtcatgcattacatgtaatt 2809
  |||||
Db 121 TCGGGACTCTAATCATATAAAACCCATCTCATATAAACGTCATGCATTACATGTTAATT 180
QY 2810 attacatgcttaacgttaaticcaacagaaattatgataatcatcgaagaccggaaca 2869
  |||||
Db 181 ATACATGCTTAAGTAATTAACAGAAATATATGATAATCATGCAAGACCGGCAACA 240
QY 2870 ggattcaatttaagaacatttattgcaaatgtttgaacgattcgtctcgagatcctcta 2929
  |||||
Db 241 GGATTCAATCTTAAGAAACTTTATTGCAAAATGTTTGAACGATCTGCTTCGGATCTCTA 300
QY 2930 gag----cgggaaagtgaattgacagatcagagatttgaagaaaaatttattacacatt 2985
  |||||
Db 301 GAGNNNCCCGAAAGTGAAATGACCGATCAGAGTTTGAAGAAAAATTTATTACACATT 360
QY 2986 tatgtaagctgaaaaaacggcctccgcaggaagcggtttttttcgttatctgattttt 3045
  |||||
Db 361 TATGTAAGCTGAAAAAAGCGGCTCGCAGGAAGCGTTTTCGTTATCTGATTTTTT 420
QY 3046 gtaaggtctgataatggttcggtgttttgaataacagccagtcgcttgagtaagaatc 3105
  |||||
Db 421 GTAAAGTCTGATAATGGTCCGTTGTTTGTAAATCAGCCAGTCGCTTGAGTAAAGAAATC 480
QY 3106 cggctcgaatttctgaagcctgagtgtatagtaataatccgcttcacgccattgttgcgcg 3165
  |||||
Db 481 CGGCTCTGAATTTCTGAAGCCTGATGTATAGTTAATATCCGTTACGCCCATGTTGTCGG 540
QY 3166 cttttcccgagggttgccttccttcgttttgaagagatgtctccgcgagtgctttccccc 3225
  |||||
Db 541 CTTTCCCGGGAGTTTGCTTCCCTGTTGAGAAAGATGCTCCGCGGATGCTTTTCCCC 600
QY 3226 ggaagcagctctgcaaggttcccttttgcacccagccgagggcttgcttctgat 3285
  |||||
Db 601 GGAGCGACGTCGCAAGGTTCCCTTTTGATGCCACCCAGCCGAGGGCTTGCTTCTGAT 660
QY 3286 ttgttaattgaattatcaggtagcttatgatgtctgaagataatccgcaaccccgcca 3345
  |||||
Db 661 TTTGTAATGTAATTAATCAGGTAGCTTATCATATGCTGAAGATAATCCGCAACCCGCTCA 720
QY 3346 aacgltgtgataaacgggtaccatgtagc 3374
  |||||
Db 721 AACGTTGTAATAACCGGTACCATTGTTGC 749

```

Search completed: November 15, 2001, 20:27:45  
 Job time: 14717 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 15, 2001, 16:18:58 ; Search time 4124.15 Seconds  
(without alignments)  
13443.009 Million cell updates/sec

Title: US-09-698-903B-1  
Perfect score: 5865  
Sequence: 1 aattacacggtatatatcc.....caattgaatatatctctgcg 5865

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues 20456230  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1:	gb_est1:	*
2:	gb_est2:	*
3:	gb_est3:	*
4:	gb_est4:	*
5:	gb_est5:	*
6:	gb_est6:	*
7:	gb_est7:	*
8:	gb_est8:	*
9:	gb_est9:	*
10:	gb_est10:	*
11:	gb_est11:	*
12:	gb_est12:	*
13:	gb_est13:	*
14:	gb_est14:	*
15:	gb_est15:	*
16:	gb_est16:	*
17:	gb_est17:	*
18:	gb_est18:	*
19:	gb_est19:	*
20:	gb_est20:	*
21:	gb_est21:	*
22:	gb_est22:	*
23:	gb_est23:	*
24:	gb_est24:	*
25:	gb_est33:	*
26:	gb_est34:	*
27:	gb_est35:	*
28:	gb_est36:	*
29:	gb_est37:	*
30:	gb_est38:	*
31:	gb_est39:	*
32:	gb_est40:	*
33:	em_estba:	*
34:	em_estfun:	*
35:	em_esthum1:	*
36:	em_esthum2:	*
37:	em_esthum3:	*
38:	em_esthum4:	*
39:	em_esthum5:	*
40:	em_esthum6:	*
41:	em_esthum7:	*
42:	em_esthum8:	*
43:	em_esthum9:	*

44:	em_esthum10:	*
45:	em_esthum11:	*
46:	em_esthum12:	*
47:	em_esthum13:	*
48:	em_esthum14:	*
49:	em_esthum15:	*
50:	em_esthum16:	*
51:	em_esthum17:	*
52:	em_esthum18:	*
53:	em_esthum19:	*
54:	em_esthum20:	*
55:	em_esthum21:	*
56:	em_esthum22:	*
57:	em_esthum23:	*
58:	em_esthum24:	*
59:	em_esthum25:	*
60:	em_esthum26:	*
61:	em_esthum27:	*
62:	em_esthum28:	*
63:	em_estin1:	*
64:	em_estin2:	*
65:	em_estin3:	*
66:	em_estin4:	*
67:	em_estin5:	*
68:	em_estom1:	*
69:	em_estom2:	*
70:	em_estov1:	*
71:	em_estov2:	*
72:	em_estpl1:	*
73:	em_estpl2:	*
74:	em_estpl3:	*
75:	em_estpl4:	*
76:	em_estpl5:	*
77:	em_estpl6:	*
78:	em_estpl7:	*
79:	em_estpl8:	*
80:	em_estpl9:	*
81:	em_estpl10:	*
82:	em_estro1:	*
83:	em_estro2:	*
84:	em_estro3:	*
85:	em_estro4:	*
86:	em_estro5:	*
87:	em_estro6:	*
88:	em_estro7:	*
89:	em_estro8:	*
90:	em_estro9:	*
91:	em_estro10:	*
92:	em_estro11:	*
93:	em_estro12:	*
94:	em_estro13:	*
95:	em_estro14:	*
96:	em_estro15:	*
97:	em_estro16:	*
98:	em_estro17:	*
99:	em_estro18:	*
100:	em_estro19:	*
101:	em_estro20:	*
102:	gb_est25:	*
103:	gb_est26:	*
104:	gb_est27:	*
105:	gb_est28:	*
106:	gb_est29:	*
107:	gb_est30:	*
108:	gb_est31:	*
109:	gb_est32:	*
110:	gb_est41:	*
111:	gb_est42:	*
112:	gb_est43:	*
113:	gb_est44:	*
114:	gb_est45:	*
115:	gb_est46:	*
116:	gb_est47:	*

117: gb\_est48:\*  
118: gb\_est49:\*  
119: gb\_est50:\*  
120: gb\_est51:\*  
121: gb\_est52:\*  
122: gb\_est53:\*  
123: gb\_est54:\*  
124: gb\_est55:\*  
125: gb\_est56:\*  
126: gb\_est57:\*  
127: gb\_est58:\*  
128: gb\_est59:\*  
129: gb\_est60:\*  
130: gb\_est61:\*  
131: gb\_est62:\*  
132: gb\_est63:\*  
133: gb\_est64:\*  
134: gb\_est65:\*  
135: gb\_est66:\*  
136: gb\_est67:\*  
137: gb\_est68:\*  
138: gb\_est69:\*  
139: gb\_est70:\*  
140: gb\_est71:\*  
141: gb\_est72:\*  
142: gb\_est73:\*  
143: gb\_est74:\*  
144: gb\_est75:\*  
145: gb\_est76:\*  
146: gb\_est77:\*  
147: gb\_est78:\*  
148: gb\_est79:\*  
149: gb\_est80:\*  
150: gb\_est81:\*  
151: gb\_est82:\*  
152: gb\_est83:\*  
153: gb\_est84:\*  
154: gb\_est85:\*  
155: gb\_est86:\*  
156: gb\_est87:\*  
157: gb\_est88:\*  
158: gb\_est89:\*  
159: gb\_est90:\*  
160: gb\_est91:\*  
161: gb\_est92:\*  
162: gb\_est93:\*  
163: gb\_est94:\*  
164: gb\_est95:\*  
165: gb\_est96:\*  
166: gb\_est97:\*  
167: gb\_est98:\*  
168: gb\_est99:\*  
169: gb\_est100:\*  
170: gb\_est101:\*  
171: gb\_est102:\*  
172: gb\_est103:\*  
173: gb\_est104:\*  
174: gb\_est105:\*  
175: gb\_est106:\*  
176: gb\_est107:\*  
177: gb\_est108:\*  
178: gb\_est109:\*  
179: gb\_est110:\*  
180: gb\_est111:\*  
181: gb\_est112:\*  
182: gb\_est113:\*  
183: gb\_est114:\*  
184: gb\_est115:\*  
185: gb\_est116:\*  
186: gb\_est117:\*  
187: gb\_est118:\*  
188: gb\_est119:\*  
189: gb\_est120:\*

190: gb\_est121:\*  
191: gb\_est122:\*  
192: gb\_est123:\*  
193: gb\_est124:\*  
194: gb\_est125:\*  
195: gb\_est126:\*  
196: gb\_est127:\*  
197: gb\_est128:\*  
198: gb\_est129:\*  
199: gb\_est130:\*  
200: gb\_est131:\*  
201: gb\_est132:\*  
202: gb\_est133:\*  
203: gb\_est134:\*  
204: gb\_est135:\*  
205: gb\_est136:\*  
206: gb\_est137:\*  
207: gb\_est138:\*  
208: gb\_est139:\*  
209: gb\_est140:\*  
210: gb\_est141:\*  
211: gb\_est142:\*  
212: gb\_est143:\*  
213: gb\_est144:\*  
214: gb\_est145:\*  
215: gb\_est146:\*  
216: gb\_est147:\*  
217: gb\_est148:\*  
218: gb\_est149:\*  
219: gb\_est150:\*  
220: gb\_est151:\*  
221: gb\_est152:\*  
222: gb\_est153:\*  
223: gb\_est154:\*  
224: gb\_est155:\*  
225: gb\_est156:\*  
226: gb\_est157:\*  
227: gb\_est158:\*  
228: gb\_est159:\*  
229: gb\_est160:\*  
230: gb\_est161:\*  
231: gb\_est162:\*  
232: gb\_est163:\*  
233: gb\_est164:\*  
234: gb\_est165:\*  
235: gb\_est166:\*  
236: gb\_est167:\*  
237: gb\_est168:\*  
238: gb\_est169:\*  
239: gb\_est170:\*  
240: gb\_est171:\*  
241: gb\_est172:\*  
242: gb\_est173:\*  
243: gb\_est174:\*  
244: gb\_est175:\*  
245: gb\_est176:\*  
246: gb\_est177:\*  
247: gb\_est178:\*  
248: gb\_est179:\*  
249: gb\_est180:\*  
250: gb\_est181:\*  
251: gb\_est182:\*  
252: gb\_est183:\*  
253: gb\_est184:\*  
254: gb\_est185:\*  
255: gb\_est186:\*  
256: gb\_est187:\*  
257: gb\_est188:\*  
258: gb\_est189:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	565.2	9.6	796	226	AQ250180	T26K17-T7
2	554.2	9.4	820	226	AQ250179	T24K7-T7
3	315.6	5.4	579	226	AQ249825	AQ249825 T26K17-T7
4	312.6	5.3	451	226	AQ249824	AQ249824 T24K7-T7
5	101	1.7	1101	219	CNS0021J	AL061936 Drosophil
6	92.6	1.6	625	220	CNS036A2	AL229763 Tetraodon
7	91.4	1.6	1101	219	CNS00EVL	AL069706 Drosophil
c 8	90.8	1.5	1101	219	CNS00EVL	AL069706 Drosophil
9	89.6	1.5	1101	219	CNS0039G	AL063921 Drosophil
10	89.2	1.5	1280	80	BF264950	BF264950 HV_CEA001
c 11	89	1.5	1101	219	CNS017V2	AL108536 Drosophil
12	88.6	1.5	942	219	CNS00601	AL065607 Drosophil
c 13	87.8	1.5	1101	219	CNS0039G	AL063921 Drosophil
c 14	87.2	1.5	1201	219	CNS0167M	AL106396 Drosophil
15	85.6	1.5	529	105	AL514657	AL514657 AL514657
16	85	1.4	633	105	AL513979	AL513979 AL513979
c 17	84.6	1.4	764	239	AZ193158	AZ193158 SP_1022_B
18	84	1.4	1151	152	BG309087	BG309087 HVSMEC000
19	83.6	1.4	1101	219	CNS00LT2	AL078714 Drosophil
20	83.4	1.4	614	219	CNS0152H	AL104915 Drosophil
21	83.4	1.4	1225	219	CNS0161D	AL106171 Drosophil
c 22	83.2	1.4	982	227	AQ325799	AQ325799 nbx0021B
c 23	83	1.4	609	220	CNS025K2	AL182171 Tetraodon
c 24	83	1.4	1101	219	CNS00LJT	AL068307 Drosophil
c 25	82.4	1.4	1032	220	CNS020TP	AL206746 Tetraodon
c 26	82	1.4	1101	219	CNS00KAE	AL07628 Drosophil
27	81.2	1.4	1106	80	BF264948	BF264948 HV_CEA001
c 28	81	1.4	821	219	CNS0090X	AL052985 Drosophil
29	80.6	1.4	945	221	CNS04D0K	AL285149 Tetraodon
30	80.4	1.4	842	219	CNS0100N	AL099305 Drosophil
c 31	80.4	1.4	945	221	CNS04D0K	AL285149 Tetraodon
32	80.4	1.4	1101	219	CNS00BO7	AL069440 Drosophil
33	80	1.4	926	80	BF264913	BF264913 HV_CEA001
c 34	80	1.4	937	219	CNS006ST	AL065880 Drosophil
35	80	1.4	1101	219	CNS003BD	AL064091 Drosophil
c 36	80	1.4	1101	219	CNS00EPO	AL069493 Drosophil
c 37	79.8	1.4	644	227	AQ327887	AQ327887 nbx0042A
c 38	79.8	1.4	934	239	AZ184244	AZ184244 SP_1002_B
39	79.6	1.4	788	219	CNS00AK8	AL055582 Drosophil
c 40	79.6	1.4	937	219	CNS0066L	AL062959 Drosophil
c 41	79.6	1.4	1059	219	CNS0022B	AL097133 Drosophil
c 42	79.2	1.4	1036	219	CNS00599	AL057797 Drosophil
c 43	79	1.3	1043	219	CNS0145P	AL103735 Drosophil
c 44	79	1.3	1101	219	CNS00EQL	AL069526 Drosophil
45	78.8	1.3	372	105	AL513971	AL513971 AL513971

ALIGNMENTS

RESULT 1	
AQ250180	
LOCUS	AQ250180 796 bp DNA GSS 06-OCT-1998
DEFINITION	T26K17-T7.1 TAMU Arabidopsis thaliana genomic clone T26K17, DNA sequence.
ACCESSION	AQ250180
VERSION	AQ250180.1
KEYWORDS	GI:3700263
SOURCE	GSS.
ORGANISM	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE	1 (bases 1 to 796)
AUTHORS	Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and Ecker,J.
TITLE	BAC End Sequences at ATGC

JOURNAL COMMENT

Unpublished (1997)  
Contact: Ecker J.  
Arabidopsis Thaliana Genome Center  
University of Pennsylvania  
Dept. of Biology, University of Pennsylvania, Philadelphia, PA 19104  
Tel: 215-898-9384  
Fax: 215-898-8780  
Email: jecker@genome.bio.upenn.edu  
Seq primer: T7  
Class: BAC ends.

FEATURES

source  
1..796  
/organism="Arabidopsis thaliana"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/clone="T26K17"  
/clone\_lib="TAMU"  
/sex="Thermaphrodite"  
/note="Vector: BelOBAcII; Site\_1: HindIII; Site\_2: HindIII"; Produced by Rod Wing"  
BASE COUNT 251 a 163 c 123 g 258 t 1 others  
ORIGIN

Query Match

Best Local Similarity 9.6%; Score 565.2; DB 226; Length 796;  
Matches 657; Conservative 0; Mismatches 99; Indels 3; Gaps 3;

Qy	964	caagcttggagtgatcggagggtctaggagatacatgagattcaagtgagcattagatctac	1023
Db	15	CATTCTTTGGAGTATCGGAGGGTCTAGGATACATGATGATCCCGTGACATGATCTAC	74
Qy	1024	accgttgagatttgagtggtgatgtgtgaggttaatttacttggttaacggccacaaa	1083
Db	75	ACCGTTGGATTTTACGTGTGGATATGTGTGACGTAAATTTACTTGTGACGGCCACAAA	134
Qy	1084	ggcctaaggagagtggtgagacccttatcgcttgacccgtggaataataagccacgtg	1143
Db	135	GGCCTAAGGAGAGTGTGTGAGACCCCTTATCGGCTTGACCCCTGGAATAATGCCACGTG	194
Qy	1144	aagtaattccatgaactcttatcgcttatgatgagtaaaattgtgagtgagtggt	1203
Db	195	AAGATAATTCCTCAATCTTATCGTATCTATGAGTGAATGTGTGATGGAGTGGT	254
Qy	1204	gcttgctcaatttacttgctggtgagacttgcccttcccttatggggaatttatattt	1263
Db	255	GCTTGCTCAATTTACTTGCCTGGTGGACTTGGCCCTTCCTTATGGGAATTTATATTT	314
Qy	1264	acttactatagagcttccatccctttttttacottggatttagttaataataatgta	1323
Db	315	ACTTACTATAGAGCTTTCATACCTTTTTCCTTGGATTTAGTAAATATAATAATGTA	374
Qy	1324	tgattcatgaataaaaatgggaaatttttgaatttgactgactgataaataagcattag	1383
Db	375	TGATTCATGAATAAAAATGGGAAATTTTGAATTTGTACTCTTAATGCATAAATAAG	434
Qy	1384	tgaactgtggaataataatttttttcaatttaaaagcaaaatttgcccttttactagaatt	1443
Db	435	TGAACCTGTGGAATATATATATTTTTCCTTTTAAACCCAAATTTGCCCTTTCTTAATTT	494
Qy	1444	ataaatagaaaaataataacattcaaaataaaaaataagaacttccaaaaaac	1503
Db	495	AT-AATTTTGAATAATTTATTTACCTTCCCAATTAATAATTTAGAACCTTCCAAAAAC	553
Qy	1504	agaactatgttaattgtgtaagatttagtcacatacaatgcatctgttacaatatgtta	1563
Db	554	CGAACCATGTTTTAAGTTTTTAAGAAATAATCCACCTCCCAATCCTCCGTACCATATGTTT	613
Qy	1564	caacaagtcaaacgcccaacaaagttagcacgtctctaaataaactaaagagtcacagaaa	1623
Db	614	ACACCAATCTTAAACCCCAACAAATTTTGC-CCTCTTAAATAACTAAAGATCC-CGAAAA	671

QY	1624	tattacaataatcaagcccaaaagttattgatcaaaaaaaacccgccaacaaagc	1683
Db	672	TTTTTCCCAATCTAAACCCCCCAAAATTTTGATCAAAAAAAACCCCCCAACCTA	731
QY	1684	taaacaaagtcacaaaaaaactctcaagtcctccatctt	1722
Db	732	CAATTCCAAAAAAACAATCCCAATTCCTCATCTCCCTTTT	770
RESULT	2		
LOCUS	AQ250179	820 bp DNA	GSS
DEFINITION	T24K7-T7.1 TAMU Arabidopsis thaliana genomic clone T24K7,		DNA
ACCESSION	AQ250179	sequence.	
VERSION	AQ250179		
KEYWORDS	AQ250179.1	GI:3700262	
SOURCE	GSS.		
ORGANISM	thale cress.		
REFERENCE	Arabidopsis thaliana		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.		
TITLE	1 (bases 1 to 820)		
JOURNAL	Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and Ecker,J.		
COMMENT	BAC End Sequences at ATCC Unpublished (1997) Contact: Ecker J. Arabidopsis Thaliana Genome Center University of Pennsylvania Dept. of Biology, University of Pennsylvania, Philadelphia, PA 19104 Tel: 215-898-9384 Fax: 215-898-8780 Email: jecker@genome.bio.upenn.edu Seq primer: T7 Class: BAC ends.		
FEATURES	Location/Qualifiers		
source	1..820		
	/organism="Arabidopsis thaliana"		
	/strain="Columbia"		
	/db_xref="taxon:3702"		
	/clone="T24K7"		
	/clone_lib="TAMU"		
	/sex="hermaphrodite"		
	/note="Vector: BeloBACII; Site_1: HindIII; Site_2: HindIII"		
	; Produced by Rod Wing"		
BASE COUNT	257 a 149 C 140 G 272 t		2 others
ORIGIN			
Query Match	9.4%; Score 554.2; DB 226; Length 820;		
Best Local Similarity	87.8%; Pred. No. 5.1e-100;		
Matches	685; Conservative 0; Mismatches 84; Indels 11; Gaps 7;		
QY	968	ctttgagtgatcgagaggtctagatcatatgagattcaagtgagctagatctacacg	1027
Db	42	CTTTGGAGTGATCGAGGGCTTAGATACATGAGATTCCACTGACTATGATCTACCG	101
QY	1028	tbgatttttgagtgaggatgtgagggttaatttacttggtaacgcccacaaagcc	1087
Db	102	TTGGATTTTGAGTGTGGATATGTGTGAGGTTAATTTACTTGGTAAACGCCAAGGCC	161
QY	1088	taagagaggtgttgagacccttatcggttgacgcgtgggaataatgccagtggaaga	1147
Db	162	TAAAGAGATGTGTGAGACCCCTTATCGCTTTGAACCGCTGGAATATGCCACGTGAAGA	221
QY	1148	taattccatgaattctatcgttacttatgagtgaaattgtgtgagtgagtgagtgctt	1207
Db	222	TAAATCCATGAATCTTATCGTTATCTATGAGTGAATTTGTGTGATGTGGAGTGTGCTT	281
QY	1208	gctcaatttacttgcctgggtggacttggcccttccttatggggaatttatatttactt	1267

BASE COUNT	203 a	79 c	60 g	232 t	5 others
ORIGIN					
Query Match	5.4%; Score 315.6; DB 226; Length 579;				
Best Local Similarity	78.7%; Pred. No. 1.4e-52;				
Matches	440; Conservative	0; Mismatches	104; Indels	15; Gaps	5;
QY	1129	aataatgccacgtggaagataatccatgaatcttctatcgttatctatgagtgaaattgtg	1188		
Db	1	AATAATGTNANGTGAATAATAATCCATGAATCTTATCGTTATCTATGAGTTGAAATGTG	60		
QY	1189	tgatggtgagtggtgcttgccttacttttacttgcotggtggaacttggcccttccctatg	1248		
Db	61	TGATGTGGAGTGTGCTCTCTCATTTTACTTTTCTGGTGGAAATTTGCCCTTTCTTATG	120		
QY	1249	gggaatttatatttactactatagagcttctacaccttttttttacccttgattttagt	1308		
Db	121	GGGAATTTATATNTTACTTACTATAGAGCTTTTCATCACCTTTTNTTAACTTGGATTTATT	180		
QY	1309	taataataatggtgatgattcatgaataaaaaatgggaatttttgaatttctgactgctaa	1368		
Db	181	TTATATTTAATGCTATGAATCATGAATAAANAATGGGAAA-TTTTGAATTTGTACTGCTAA	239		
QY	1369	atgcataagcttaggtgaaactgtggaatatatatatttttctcatttaaaagcaaaattg	1428		
Db	240	ATGCATAAAATTTAGGTGAAACTGTGGAATATATATTTTTTTCATTTTAAAGCAAAATTTG	299		
QY	1429	ccctttactagaattacaatatagaaaaatatataacattcaaaataaaaaatgaaaaataa	1488		
Db	300	CCTTTTACTAAAATTTTAAATTTAGAAAAAT-TTTTACATTTCCAAATTTTAAATGAAATAA	358		
QY	1489	gaacttcaaaaaacagaaactgttttaattgtgtaaagattagtcgcacatcaagtcatc	1548		
Db	359	GACTTTTCAAAAACC--AAATATTTTAAATTTTAAATAA-----ACCCCAAAATCAAT	409		
QY	1549	tgtacaatatgtttacaacaagtcataagcccaacaagaagtagcacgctcaataaacta	1608		
Db	410	TTTTCTTTTAAAAAATTTTAAAAAATTTTAAACCCCAAAAATTTTCCCTTCTCAATAATTA	469		
QY	1609	aagagtcacgaaaatattacaatatgaagcccaacaaagattattgatcaaaaaaaaaa	1668		
Db	470	AA----TTCCGAAAAATTTTCAATCTTTTACCCCAAAAATTTTTTTATTGAAAAA	525		
QY	1669	aacgcccacaaagctaaa	1687		
Db	526	ACCCCTCAAAATTAATAAA	544		
RESULT	4				
AQ249824					
LOCUS	T24K7-T7 TAMU Arabidopsis thaliana genomic clone T24K7, DNA				
DEFINITION	sequence.				
ACCESSION	AQ249824				
VERSION	AQ249824.1	GI:3699907			
KEYWORDS	GSS.				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 451)				
AUTHORS	Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and Ecker, J.				
TITLE	BAC End Sequences at ATGC				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Ecker J. Arabidopsis Thaliana Genome Center University of Pennsylvania Dept. of Biology, University of Pennsylvania, Philadelphia, PA 19104 Tel: 215-898-9384				

Fax: 215-898-8780				
Email: jecker@atgenome.bio.upenn.edu				
Seq primer: T7				
Class: BAC ends.				
FEATURES				
source	1..451			
	/organism="Arabidopsis thaliana"			
	/strain="Columbia"			
	/db_xref="taxon:3702"			
	/clone="T24K7"			
	/clone_lib="TAMU"			
	/sex="hermaphrodite"			
	/note="Vector: BelOBACII; Site_1: HindIII; Site_2: HindIII"			
BASE COUNT	145 a 54 c 64 g 185 t 3 others			
ORIGIN				
Query Match	5.3%; Score 312.6; DB 226; Length 451;			
Best Local Similarity	84.6%; Pred. No. 5.3e-52;			
Matches	384; Conservative 0; Mismatches 67; Indels 3; Gaps 3;			
QY	1118	tgaaacctggaataatgccacgtggaagataattccatgaactcttctgttctatctatga	1177	
Db	1	TGAACCGCTGGGAATAATGCCACGTGGAATATAATTCATGAATCTTATCGTTATCTATGA	60	
QY	1178	gtgaattgtgtagtggagtggtgcttgcctatttctacttgcctggtagacttgcc	1237	
Db	61	GTGANATTGTGTGATGGTGGAGTGGTGTGCTTCATTTTACTTTGCTGGTGGACTTGCC	120	
QY	1238	cttccctttaggggaatttatattttacttactatagagcttctcacaaccttttttacc	1297	
Db	121	CTTTCCTTATGGGAATTTATATTTTACTTACTATAAAAGTTTTCAAAACCTTTTNTNTAAC	180	
QY	1298	ttagatttagtaataataataatggtagtatctatcaataaaaaatgggaatttttgaatt	1357	
Db	181	TTTGTATTTTTTTTTTATATATGGTATGAATCATGAATGAACATGCCATCCITGGAATT	240	
QY	1358	tgtactgctaaaatgcataagattagtgaaactgtggaactatataatttttctcatttaa	1417	
Db	241	TGTACTGCTTAATGCATAAAAATAGTGAAAA-TGTGGAATATATA-TTTTTCATTATAA	298	
QY	1418	agcaaatgttgccttttactagaattataataatagaaaaataataaacatcaataaa	1477	
Db	299	AGCAAAATTTGCCCTTTTATTAATAATTTAGAAAAAT-TTTTACTTTTCCAATAAA	357	
QY	1478	aatgaataaagaactttcaaaaaacagaactatgtttaatgtgaaagattagtcgcac	1537	
Db	358	ATTGAAAAATTAACACTTTTCAAAAACAGAACTATGTGTTATTGTATATAAAATATTCCAA	417	
QY	1538	atcaagtcatcttaccataatgtttacaacaagt	1571	
Db	418	TCTAGTTTCATCTGTTTAATAATTTTATAACAATT	451	
RESULT	5			
CNS0021J				
LOCUS	CNS0021J	1101 bp	DNA	GSS
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR05N11 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.			
ACCESSION	AL061936			
VERSION	AL061936.1	GI:4940214		
KEYWORDS	GSS.			
SOURCE	fruit fly.			
ORGANISM	Drosophila melanogaster			
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
AUTHORS	1 (bases 1 to 1101)			
TITLE	Genoscope.			
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage			

Fax: 215-898-8780	
Email: jecker@atgenome.bio.upenn.edu	
Seq primer: T7	
Class: BAC ends.	
FEATURES	Location/Qualifiers
source	1..451
	/organism="Arabidopsis thaliana"
	/strain="Columbia"
	/db_xref="taxon:3702"
	/clone="T24K7"
	/clone_lib="TAMU"
	/sex="hermaphrodite"
	/note="Vector: BelobACII; Site_1: HindIII; Site_2: HindIII"
	; Produced by Rod Wing"
BASE COUNT	145 a 54 c 64 g 185 t 3 others
ORIGIN	
Query Match	5.3%; Score 312.6; DB 226; Length 451;
Best Local Similarity	84.6%; Pred. No. 5.3e-52;
Matches 384; Conservative	0; Mismatches 67; Indels 3; Gaps 3;
QY	1118 tgaaccgctgaataatgccacgtggaagataattccatgaatcttcttcttatctatga 1177
Db	1 TGAACCCGCTGGAAATAATGCCACGTGGAATATAATTCATGAATCTTATPCGTTATCTATGA 60
QY	1178 gtgaaattgtgtgatggtggtgagtggtctgctcatcttttacttgcctggcggaacttggcc 1237
Db	61 GTGANATTTGTGATGCTGGAGTGGTCTTCTCATTTTACTTGCCTGGTGGACTTGGCC 120
QY	1238 ctttccttatgggaatttatattttacttacttagagcttctcatccttttttttacc 1297
Db	121 CTTTCCCTATGGGAATTTATATTTTACTTACTATAAAGTTTTCAAAACCTTTTNTNTAAC 180
QY	1298 ttgattactgaataataatgggtatgattcatgataaaaaatggaaaatttttgaatt 1357
Db	181 TTTGTATTTTTTTTTTATATGTTATGTTATGAATCATGAATAACAATGGGACATCTTGAATT 240
QY	1358 tgactgctaaatgcataaagattagtgtaaactgtgaaactatggaatatattttttcttataa 1417
Db	241 TGFACTCTAAATGCATAAANAATAGGTGAAA-TGTGGAATATATA-TTTTTTCATTATAA 298
QY	1418 agcaaatgttccttttacttagaattataataatagaaaaatatatacaattcacaataaa 1477
Db	299 AGCAAAATTTGCCCTTTTATTAATAATTTAATTTAGAAAAAT-TTTTACTTTCCAATAA 357
QY	1478 aatgaaaaataagaactttcaaaaaacagaaactatgtttaatgtgtaagaagattagtcgac 1537
Db	358 ATTGAAAATTAACCTTCTCAAAAACAGAACTATGTTTATTTTATAAAAATTTATCCAA 417
QY	1538 atcaagtcactgttacaatatgtttacaacaagt 1571
Db	418 TCTAGTTTCATCTGTTTAAATATTTTATAACAATT 451
RESULT	5
CNS0021J	
LOCUS	1101 bp DNA GSS 03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR05N11 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL061936
VERSION	AL061936.1 GI:4940214
KEYWORDS	GSS.
SOURCE	fruit fly.
ORGANISM	Drosophila melanogaster
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyaroida; Drosophilidae; Drosophila.
AUTHORS	1 (bases 1 to 1101)
TITLE	Genoscope.
JOURNAL	Direct Submission
	Submitted (02-JUN-1999); Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

**COMMENT**  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RpCi-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

# FEATURES

source  
1. 1101  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone\_lib="RPCi-98"  
/clone="BACR05N11"  
/note="End : TET3"  
BASE COUNT 631 a 7 c 28 g 289 t 146 others  
ORIGIN

Query Match 1.7%; Score 101; DB 219; Length 1101;  
Best Local Similarity 40.0%; Pred. No. 6,6e-10;  
Matches 346; Conservative 41; Mismatches 479; Indels 0; Gaps 0;

QY 1254 tttatttttactatagaggtttcattacaccttttttaoctggattgatta 1313  
Db 13 TTTTCKTNTTNNTTGGTTTNNNTTNNCTCTTCTTNTTNNNTKTTTNTTTA 72

QY 1314 tataatggtatgttcataaataatggaaattttgaaattgtactgctaaatgca 1373  
Db 73 TTTTNNNAANAANKAAATNATAATTTTAAATTAATAAKAKDFAAATAAANAHA 132

QY 1374 taagattagtgaaactgtggaatatatttttccatttaaaagcaaatgtcctt 1433  
Db 133 AAAAAGTGTTCATGAAAGTTTAAAAAAAATAAATAAATAAATAAATAAATAA 192

QY 1434 tactagaattataatagaaaataatacaattcaataaataaataaataaact 1493  
Db 193 NNAANNAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 252

QY 1494 tccaataaacagaaactatgtttatgtgtaagattagtcgcacatcaagtcctgtta 1553  
Db 253 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 312

QY 1554 caatatgtttacaagaagtcagcccaagaagtagcagctctaaataaactaaagag 1613  
Db 313 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 372

QY 1614 tcacgaaataattacaaatcaaaagcccaagaagttattgatacaaaaaaaagcgc 1673  
Db 373 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 432

QY 1674 ccacaaagctaaacaaagtcacaaataaactctcaagtcctccatttgaac 1733  
Db 433 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 492

QY 1734 attgaaactatcacaaacagtcagataaatctctctctggcctgtctcccaacc 1793  
Db 493 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 552

QY 1794 tctacataactccctacgcgattgaattgtttactgtacacctttccggtgcaatgat 1853  
Db 553 TTTTNTTAAATTTTAAATTTTAAAWAAATTTAAATAAANAATTTTAAATTTTAA 612

QY 1854 attgattagttgttgaaactaatagggttgtaacaatcgaagtcaggaatggatt 1913

Db 613 AWWAAAAAAAWTTTTTAAAWTWTTTTTTTTWTAAATTAATAWAAAAAAWATAWAAWTT 672  
QY 1914 tggccaagattttcccgagagcttctctagtagaagaagcccatccagagaaatttactagta 1973  
Db 673 WTTTATWATWAAATWAAATTTTAAAAWTTTTTTTAAWTTTTTTTWTAAATTTTATTA 732

QY 1974 aaataaatacccaattgggtttcttattatgtgccaaattcaataataattatagaggata 2033  
Db 733 WTTTAAATAAATWATWWTAAATTTTAAATTTTAAWATAWATAWATAWATAWATAWATA 792

QY 2034 ttccaatgaaacgtagatgaattattagtaaatagtcaggtgaggaacatttaaaaaatc 2093  
Db 793 TTTTNTTAAATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 852

QY 2094 ctacgtcagatattcaactttcaaaa 2119  
Db 853 TTTTNTTAAATTTATTTTWTATTATAA 878

## RESULT 6

CNS036A2 625 bp DNA GSS 15-MAY-2000  
LOCUS Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone  
DEFINITION 215D15 of library G from Tetraodon nigroviridis, genomic survey  
sequence.

ACCESSION AL229763

VERSION AL229763.1 GI:7888758

KEYWORDS GSS; genome survey sequence.

SOURCE Tetraodon nigroviridis.

ORGANISM Tetraodon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontidae; Tetraodon.

REFERENCE 1 (bases 1 to 625)

AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,  
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and  
Weissenbach,J.

TITLE Characterization and repeat analysis of the compact genome of the  
freshwater pufferfish Tetraodon nigroviridis

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 625)

AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,  
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,  
Saurin,W. and Weissenbach,J.

TITLE Human gene number estimate provided by genome wide analysis using  
Tetraodon nigroviridis DNA sequence

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 625)

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases

COMMENT This sequence is a single read and was generated as part of a large  
scale clone-end sequencing project of the Tetraodon nigroviridis  
genome. For more information, please take a look at  
<http://www.genoscope.cns.fr/Tetraodon>.

## FEATURES

source

1. 625  
Location/Qualifiers

/organism="Tetraodon nigroviridis"

/db\_xref="taxon:99883"

/clone="215D15"

/clone\_lib="G"

/note="Genoscope sequence ID : C0AG215CB08SP1-end ;  
PUC-Ori"

BASE COUNT 394 a 39 c 4 g 170 t 18 others

## ORIGIN

Query Match 1.6%; Score 92.6; DB 220; Length 625;  
Best Local Similarity 49.0%; Pred. No. 3e-08;  
Matches 226; Conservative 6; Mismatches 231; Indels 0; Gaps 0;



```
QY 1239 tttctctatgggaattattattttactactatagctgttttcataccttttttttaccct 1298
Db 136 TTTTCTCTATGGGAATTATTATTACTACTATAGCTGTTTTCATACTCTTTTCTTACCCT 195
QY 1299 tggatttagtataataatggtatgatcatgaataaaatgggaatttttgaattt 1358
Db 196 TTTTCTCTATGGGAATTATTATTACTACTATAGCTGTTTTCATACTCTTTTCTTACCCT 255
QY 1359 gtactgctaaatgataagattagggtgaactgtggaaatataatatttttttcattttaa 1418
Db 256 ATAAWATATATATAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 315
QY 1419 gcaaaatttgccttttctactagaattataataatagaaaaataatacattcaataaaa 1478
Db 316 TWAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 375
QY 1479 atgaataaagaactttcacaacagaaactatgttttaagtgtgaagatttagcgcaca 1538
Db 376 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 435
QY 1539 tcaagtcatctgtttacataatgtttcacacaagtcataagcccaacaaagtttagcagct 1598
Db 436 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 495
QY 1599 aaataaactaaagagtcacagaaaattattacaatctatagcccaacaaagtttagtc 1658
Db 496 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 555
QY 1659 aaaaaaataaagcccaacaaagtcataacaaagtcacaaataa 1703
Db 556 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 600

RESULT 7
CNS00EVL 1101 bp DNA GSS 04-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC:
DEFINITION BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL069706
VERSION AL069706.1 GI:4949849
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL

COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osogawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.
FEATURES
source 1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
```

```
BASE COUNT 419 a 91 c 60 g 299 t 232 others
ORIGIN

Query Match 1.6%; Score 91.4; DB 219; Length 1101;
Best Local Similarity 34.5%; Pred. No. 5.4e-08;
Matches 206; Conservative 100; Mismatches 291; Indels 0; Gaps 0;

QY 4275 ttttttattggtccaaataagtgagtttttttagatttcaaaaatgatttaattttttt 4334
Db 504 WWWWATTTTWWWWTWATTTTWWWWTTTAAATTAATAAATAAATAAATAAATAAATAAATAA 563
QY 4335 ttactacagtcgacctggagtaaatggctgttgaggatgtgttagaaatgtttatgtgaa 4394
Db 564 ATTAATAAATTTAAAWWWTATATTAATTAATAAATTTAAATAAATAAATAAATAAATTT 623
QY 4395 gaaatagtaaggttcaatgatcaatttcattctgctatttaattgttaaaatgtgaatttc 4454
Db 624 WATAAATAATTTTAAATAAATTAATTAATTAATTAATAAATAAATAAATAAATAAATAA 683
QY 4455 ttaactctgtgtgaacacacacacacacacacacacacacacacacacacacacacacac 4514
Db 684 WAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 743
QY 4515 atatttgggaagcgactaaataaaacttttctcatattatcacgaacctaaacacagca 4574
Db 744 AATWAWATAATATWATAATATATTTTAWWWATWWAWWWTATATAWATAWATAWAAWAA 803
QY 4575 tatggtgatttcttagggaaatcaataactcaataaatttaataaagaagaacaaagtatcaa 4634
Db 804 WATAAATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATA 863
QY 4635 tacatatgatttacacgcgtcaaacacacgaatttcgataatttaataataaagaatta 4694
Db 864 WAATWATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 923
QY 4695 atccaaatagcctccacccctataacttaactaaataaaccagcgaatgtattata 4754
Db 924 AAAAAATAAAWWWWTWATATTTTATTAATAATTTATTTWATTTWATTTWATTTWATTTW 983
QY 4755 tgcataatttatattataatctgtatataatctgtatataatctgtatataatctgtat 4814
Db 984 TATWTATATATWWTAWATAATATTTTAAWWTATATTTTAAWWTATATTTTAAWWTATATA 1043
QY 4815 tggtagaaaaagtaaacaaattataatagccgctatttggtagaaatcccttaata 4871
Db 1044 WWTWATATATAWAAWAAATTAATTTATATATATAWATAWATAWATAWATAWATAWATA 1100

RESULT 8
CNS00EVL/c 1101 bp DNA GSS 04-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC:
DEFINITION BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL069706
VERSION AL069706.1 GI:4949849
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL

COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
- Web : www.genoscope.cns.fr
Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
```

The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.fruitfly.org>. The BDGP *Drosophila melanogaster* BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed in Buffalo, ECORI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw, the same strain used for the BDGP's p1 and Est libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES source

```

t name=JG source
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_1_id="RpCI-98"
/clone_2_id="BACR3B23"
/note="end : T7"

BASE COUNT 419 a 51 c 60 g 299 t 232 others
ORIGIN

```

Query Match	1.5%	Score 90.8;	DB 219;	Length 1101;
Best Local Similarity	35.7%;	Pred. No. 7e-08;		
Matches 218;	Conservative 109;	Mismatches 276;	Indels 7;	Gaps 2;

  

QY	2012	tccaataaatattatagaaggaatttcacaaatgaaaaacgatgaattgttataggtaaaagtct	2071
Db	1100	TWWTATWTTTWTWTWTWTWTATATATAAATAAAATATWWTATATATWTAATWTTAT	1041
QY	2072	caggtaagacataaaaaaaaa--accctacgtcagatatcaacctttcaaaaatcgatcag	2128
Db	1040	WATATATTAWTTTWTAAAAATATAWWTTAAATAATATATWTAATWTAATATAWATAWAA	981
QY	2129	tgtgggaattgcacaaaaatttgggactactatataataatgccttcacaacacttgg	2188
Db	980	WTATAWATTWAAATWATAWAWTTTAAATAAAATATWAAWTTTATTTTTTTTTT	921
QY	2189	atttttttttgaggccgggaattttaactcacatatttgtttggccatgcaccaactc	2248
Db	920	TTTTTTTWTATWTTTATATWTTWAAAANWAAWTTATTTTTTWTATWATTWATA	861
QY	2249	attgttttagtgtaactccttgatttggctgcacaaatcatggttcgtgatatatttgttaa	2308
Db	860	TTAAATTTTTTWTATWTTATATATATWATWTTATWTTWTTWTTWATWTTATTTATWTT	801
QY	2309	aattctcttgaccatacacacacacatatatatatatatatatatatatatatcatcgt	2368
Db	800	WTTWTATWTTATWTTAAWTTWTTWATWTTWTTAAWTTATATATWTTATTTWTTWTTT	741
QY	2369	cactttcaattgaaaaaataatatatatatatagtcgattttcttcaaacacatata	2428
Db	740	WTTTATTTWTTTAAWTTATWTTATWTTATWTTWTTWTTTTTTTTTTTTTTWTATWAAA	681
QY	2429	tgttcgcatgacgtcgaacaaataactgcagagtaatgaaaaatataatctattgctgaa	2488
Db	680	TTATTTAAATTAAWNAATWTTATTTTAAATAAATAATTAATTTAAATTTTATWAAA	621
QY	2489	----attatctcagatgttcaagattttctottaagtcaaatctcttcaaatttttagctaaa	2544
Db	620	AAATATTTTTTTTATATTAATWAAWTTTATWATTAATATAWTTTWTAAATTTTATWTA	561
QY	2545	gtctgtcaactcaactgaagaataatacacacaaatctcgaccacggaaaaaacacataaa	2604
Db	560	TTWWWTTAAATATWTTTTTTTTTTATWTAWAAWTTWAAWAAATWAAWWAAAAATWMMTW	501
QY	2605	atttgaattt 2614	
Db	500	TTTTTTWATTW 491	

## RESULT 9

RESULT	9	
CNSG0039G		
LOCUS	1101 bp	DNA
DEFINITION	Drosophila melanogaster genome survey sequence TPT3 end of BAC # BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	GSS
ACCESSION	AL063921	
VERSION	AL063921.1	GI:4941778
KEYWORDS	GSS.	
SOURCE	fruit fly.	
ORGANISM	Drosophila melanogaster	
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
AUTHORS	1 (bases 1 to 1101)	
TITLE	Genoscope.	
JOURNAL	Direct Submission	
COMMENT	<p>Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)</p> <p>- Web : <a href="http://www.genoscope.cns.fr">www.genoscope.cns.fr</a></p> <p>Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <a href="http://www.fruitfly.org">http://www.fruitfly.org</a> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <a href="http://bacpac.med.buffalo.edu/drosophila_bac.htm">http://bacpac.med.buffalo.edu/drosophila_bac.htm</a>.</p>	

FEATURES source

```

Source
0.001101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCL-98"
/clone="BACR08K10"
/note="end : TET3"
BASE COUNT      201 a      64 c    131 g    202 t    503 others
ORIGIN

Query Match          1.5%; Score 89.6; DB 219; Length 1101;
Best Local Similarity 18.0%; Pred. No. 1.2e-07;
Matches 130; Conservative 300; Mismatches 294; Indels 0; Gaps 0;

QY 1377 gattagtgaaactgtggaatatatatttttccatttaaaggcgaataattgccttttc 1436
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 377 GRGGRGGGGGGRGGGGMATATAWMMWTTTTTTTTANAAAWAAAATAATTWAAWA 436

QY 1437 tagaatataaatatagaaaaatatataaccattcaataaaaagaataagaacttcc 1496
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 437 WAAAAAATTWAAAAWAAAANAATWMTTAAWWAAAAAATAATTTTTTTTTTTTT 496

QY 1497 aaaaaacagaacatatgtttaatgttaagattagtcacatcacagtcatctgttcaa 1556
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 497 AWTTWATAWTTTWWWTTAAAAWAAAAAAAAAAAAWAAAATAATTTWWTTTWTTW 556

QY 1557 tatgttacacaagtcataagcccacaaagtttagcagctctaataaacataagagctcc 1616
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 557 AAATAAACAMCAAWYYHHYTYYTHYYTYTWYTWYTHWYTHWYTHWYTHWYTHWYTHYAM 616

QY 1617 acgaaataattacaatcataagccccacaaagtattgatccaataaaaaaaaaaacgccca 1676
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 617 WHMTWHHTMVAWMMHWTYYTAAYVYUYTCMYYYHYMHMHHAHAHAANAATTTHTWTHYAH 676

QY 1677 acaagactaaacaagtcacaaaaaaaccttccaagctccaagctctccattcttatgaacatt 1736
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 677 WATYHYYYMYCAMCMCTHTCHCYYYYHHHTAHHTHTHHWYAHYNYMYWYMYWYMYC 736

```

[illegible]

[illegible]

isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES	source	Location/Qualifiers
BASE COUNT	404 a	54 c 40 g 251 t 193 others
ORIGIN		
Query Match	1.5%	Score 88.6; DB 219; Length 942;
Best Local Similarity	40.5%	Pred. No. 1.9e-07;
Matches	238; Conservative	46; Mismatches 304; Indels 0; Gaps 0;
QY 1169	tatctatgagtgaaatcgtgtgatggtggagcgggtcgtctgcctcattttcactgtcctcggtg	1228
DB 311	TWTATATTT	370
QY 1229	gacttgcccttctctatggggaatttatatttacttactatagacttctacacct	1288
DB 371	TT	430
QY 1289	tttttaccctggatttagttaataataatcgtgtatgatcatgaataaaatgggaat	1348
DB 431	TT	490
QY 1349	ttttgaatttgactgctaaatgcataagattagtggaactgtggaatatatatattttt	1408
DB 491	TNGTGTGTTTTTTTTTTTTTTGGGGTGTGGGGGTGTTTTTTTGTGTGTTTGTG	550
QY 1409	tcattttaaagcaaaatttgccttttactagaattataaatatagaaaaataataacat	1468
DB 551	TTGTGAARAAAAAATAAAAAAATAAAAAAAGRWAAARAKAAAAAATAAAAAA	610
QY 1469	tcaataataaataagaaataagaactttcaaaaaacagaaactatgttaatgtgtaagat	1528
DB 611	AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	670
QY 1529	tagtcgcacatcaagtcctctgttacaatatgttacaacaagtcataagccacaagat	1588
DB 671	AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	730
QY 1589	tagcacgtctaaataaactaaagagtcacagaaatattacaataatcataagcccaaaaa	1648
DB 731	AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	790
QY 1649	gtatttgatcaaaaaaaacacgcccacaagaagtcacaagaagtcctcaaaaaaaactct	1708
DB 791	AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	850
QY 1709	caagctccatctctcttattgaacattgaaactatacacaaacaa	1756
DB 851	AMAMAMAMCMMMMACMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM	898
RESULT 13		
CNS0039G/c		
LOCUS	1101 bp	DNA
DEFINITION	CNS0039G	GSS
	BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	03-JUN-1999
ACCESSION	AL063921	
VERSION	AL063921.1	GI:4941778
KEYWORDS	GSS.	
SOURCE	fruit fly.	
ORGANISM	Drosophila melanogaster	
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1101)  
Genoscope.  
Direct Submission  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila  
melanogaster genome using these BACs. For further information  
please see <http://www.fruitfly.org> The BDGP Drosophila  
melanogaster BAC library was prepared by Kazutoyo Osoeawa and  
Aaron Mammeter in Piter de Jong's laboratory in the Department of  
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
NY. The library is named RPCI-98 and was constructed by partial  
EcoRI digestion of Drosophila DNA provided by the BDGP from the  
isogenic strain y2; cn bw sp, the same strain used for the BDGP's  
p1 and EST libraries. A more detailed description of the library  
and how to order individual BAC clones, the entire library, or  
filters for hybridization from the BACPAC Resource Center can be  
found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

[illegible]

Qy	4669	taaatatttaataataagaagaaattaacccaagttagctccaccctcataaacctaacta	4728
Dd	:	: : : :	:
Dd	563	TATATTTAARAANAAWAWAATAATTATTTTWTWWTTTTTT- - --TTTWTTWTTAAWWW	509
Qy	4729	aacaataacegggaagtgttatatatcatgcatacaatttataattataatgigtataaatc	4788
Dd	:	: : : :	:
Dd	508	AANNATNAWAATAAAAANAANAATTTTTTTTTTTTTTTTGAATTAANWTATWTWTWT	449
Qy	4789	tataatacaaagtataatctctgatgttcatttttgaggaaaaagtaaacaataataatagcccgc	4848
Dd	:	: : : :	:
Dd	448	TWAAATTTTTTTTWTWWTAAATTTTWTWWTAAAAAAMMWMTWTATATKCCCCC	389
Qy	4849	t 4849	
Dd	:	:	:
Dd	388	y 388	
RESULT 14			
CNSO167M/C	LOCUS	1201 bp DNA GSS	26-JUL-1999
DEFINITION	Drosophila melanogaster genome survey sequence T7 end Of BAC BACN15M24 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
ACCSSION	AL106396	GI:5621701	
VERSION	GSS:		
KEYWORDS	fruit fly.		
SOURCE	Plasmid Drosophila melanogaster		
ORGANISM	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyridioidea; Drosophilidae; Drosophila. 1 (bases 1 to 1201)		
REFERENCE	Genoscope.		
AUTHORS	Direct Submission		
TITLE	Submitted (23-JUL-1999); Genoscope - Centre National de Sequencage		
JOURNAL	BP 191 91006 EVRY Cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr) Web : www.genoscope.cns.fr		
COMMENT	Determination of this BAC-end sequence was carried out as part of collaboration with the European Drosophila Genome Project (EDGP) http://www.edgp.ebi.ac.uk . This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Buchet and Genevieve Payan. It has been constructed in the vector pBLOBACL1.		
FEATURES	Location/Qualifiers		
source	1..1201		
	/organism="Drosophila melanogaster"		
	/plasmid="pblobacl1"		
	/db_xref="taxon:7227"		
	/clone_lib="DrosBAC"		
	/clone="BACN15M24"		
	/note="end : T7"		
BASE COUNT	323 a	87 c	79 g    551 t    161 others
ORIGIN			
Query Match	1.5%; Score 87.2; DB 219; Length 1201;		
Best Local Similarity	41.4%; Pred. No. 3.7e-07;		
Matches 205;	Conservative 65; Mismatches 217; Indels 8; Caps		
Qy	1261	ttacctactatagagcgtttcacaccttttttttacctcgatttagttaataataatg	1320
Dd	:	: : : :	:
Dd	1058	TTTWTWNATATAWAATTTTTTTTTTTTTTTTATANAANAANAANAATTTTAAA	999
Qy	1321	gtatgatctgaataaaaatgggaatttttgtactgtactgaaatgcataagatt	1380
Dd	:	: : : :	:
Dd	998	AATAAATTAATATWAAAAATTTTAAAAATTTTWTATWWTTTTTAAAAAANAATA	939
Qy	1381	eggfgaaactgtgga-----atatatattttttcoatttaaagcaaaaattgcc	1435
Dd	:	: : : :	:
Dd	938	TWAAATTTTTTTTATATWATAAANAATTTTTTTTAAAAAANAATTTAAAAATTA	879



**THIS PAGE BLANK (USPTO)**